CFTRI MYSORE
388
Principles of nu

368 ① human body
② proximate constituents
③ biomolecules
④ enzymes ⑤ energy
⑥ tissues ⑦ muscles ⑧ nerves
⑨ central nervous system
⑩ chemical senses
⑪ sense organs
⑫ blood ⑬ circulation
⑭ respiration ⑮ human physiology
ex
⑲ body temperature regulation
⑳ endocrine glands
㉑ human reproduction.

TLM

PRINCIPLES OF HUMAN PHYSIOLOGY

RECENT ADVANCES IN PHYSIOLOGY

Seventh Edition. By W. H. Newton, M.D., D.Sc., 90 Illustrations. *About* 21*s*.

RECENT ADVANCES IN SEX AND REPRODUCTIVE PHYSIOLOGY

By J. M. Robson, M.D., D.Sc., F.R.S.(Edin). *Third Edition.* 21*s*.

PHYSIOLOGY OF THE EYE

By Hugh Davson, D.Sc. 301 Illustrations. 32*s*.

PHYSICAL CHEMISTRY OF CELLS AND TISSUES

By Rudolph Höber, M.D., and Collaborators. 70 Illustrations. 42*s*.

PRACTICAL PHYSIOLOGICAL CHEMISTRY

By P. B. Hawk, Ph.D., Bernard L. Oser, Ph.D., and W. H. Summerson, Ph.D. *Twelfth Edition.* 329 Illustrations with 5 Coloured Plates. 50*s*.

EXPERIMENTAL PHYSIOLOGY for Medical Students

By D. T. Harris, M.D., D.Sc. *Fourth Edition.* 257 Illustrations. 18*s*.

BIOCHEMISTRY FOR MEDICAL STUDENTS

By W. V. Thorpe, M.A., Ph.D. *Fourth Edition.* 36 Illustrations. 18*s*.

A TEXTBOOK OF BIOCHEMISTRY : For Students of Medicine and Science

By A. T. Cameron, *C.M.G.*, D.Sc., F.R.I.C., F.R.S.C. *Sixth Edition.* 3 Plates and 25 Text-figures. 18*s*.

MICRO-ANALYSIS IN MEDICAL BIOCHEMISTRY

By E. J. King, M.A., Ph.D. 10 Illustrations. 10*s*. 6*d*.

J. & A. CHURCHILL LTD.

CFTRI-MYSORE
368
Principles of hu..

BOOK IV

The Supplying of Information—The Special Senses

Sense organs

(By H. Hartridge, F.R.S.)

CHAPTER PAGE

XXIV. THE NERVOUS MECHANISM OF SENSATION 367
XXV. VISION 372
XXVI. HEARING AND SPEECH 450
XXVII. CUTANEOUS SENSATIONS 474
XXVIII. SENSATIONS OF SMELL AND TASTE 483

BOOK V

Systems for Distribution of Materials—Blood, Circulation and Respiration

Historical Note 488
XXIX. THE BLOOD 493
XXX. GENERAL FEATURES OF THE CIRCULATION 537
XXXI. PHYSIOLOGY OF THE HEART 541
XXXII. FACTORS WHICH MODIFY THE HEART'S ACTION . . . 571
XXXIII. DYNAMICS OF THE CIRCULATION 609
XXXIV. LOCAL REGULATION OF BLOOD FLOW 642
XXXV. THE PULMONARY, HEPATIC AND CEREBRAL CIRCULATIONS 665
XXXVI. EXAMPLES OF CIRCULATORY ADAPTATION 673
XXXVII. LYMPH, TISSUE FLUIDS, AND CEREBRO-SPINAL FLUID . 685
XXXVIII. THE DEFENCE OF THE ORGANISM AGAINST INFECTION . 697
XXXIX. RESPIRATION. (I) MECHANICS OF RESPIRATORY MOVEMENTS 708
(II) THE CHEMISTRY OF RESPIRATION . . 722
XL. THE REGULATION OF THE RESPIRATORY MOVEMENTS . 749

BOOK VI

The Intake of Materials—Nutrition, Metabolism

Historical Note 781
XLI. THE EXCHANGES OF MATTER AND ENERGY IN THE BODY 784
XLII. THE VITAMINS 819
XLIII. THE NORMAL DIET OF MAN 832
XLIV. THE PHYSIOLOGY OF DIGESTION—CHANGES UNDERGONE BY THE FOODSTUFFS IN THE ALIMENTARY CANAL . 842
XLV. THE ABSORPTION OF THE FOODSTUFFS 903
XLVI. THE HISTORY OF THE FOODSTUFFS 920

CONTENTS

CHAPTER PAGE
I. INTRODUCTION 1

BOOK I

General Principles

A. Structural Principles

II. THE STRUCTURAL BASIS OF THE BODY 9

B. Biochemical Principles

III. THE ELEMENTARY CONSTITUENTS OF PROTOPLASM . . 22
IV. THE PROXIMATE CONSTITUENTS 25

C. Biophysical Principles

V. THE ENERGY OF MOLECULES IN SOLUTION . . . 54
VI. THE PROPERTIES OF COLLOIDS 69
VII. THE PASSAGE OF WATER AND SOLUTES THROUGH MEMBRANES 79
VIII. BIOELECTRIC POTENTIALS 85
IX. ENZYMES AND ENZYME ACTION 88

BOOK II

Tissues Subserving Movement and Conduction—Muscle and Nerve

Historical Note 103
X. VOLUNTARY STRIATED MUSCLE 105
XI. PLAIN MUSCLE AND OTHER CONTRACTILE TISSUES . . 162
XII. NERVE FIBRES 171

BOOK III

Centralized Control and Co-ordination—The Central Nervous System

Historical Note 198
XIII. GENERAL FEATURES OF A CENTRAL NERVOUS SYSTEM . 201
XIV. THE RECEPTORS 216
XV. THE SPINAL CORD AND ITS RELATIONSHIP TO THE HIGHER PARTS OF THE CENTRAL NERVOUS SYSTEM . . . 222
XVI. THE EXCITABILITY OF NEURONES AND THE PROPERTIES OF CONDUCTION AT SYNAPSES 236
XVII. CHARACTERISTICS OF REFLEX ACTIONS 250
XVIII. THE LABYRINTH, AND POSTURAL REFLEXES . . . 267
XIX. HIGHER AFFERENT PATHS 284
XX. THE CEREBRAL HEMISPHERES AND THE EFFERENT PATHS 293
XXI. FUNCTIONS OF THE CEREBRAL CORTEX 318
XXII. THE CEREBELLUM. 332
XXIII. THE AUTONOMIC SYSTEM: THE HYPOTHALAMUS: HUMORAL TRANSMISSION 343

PREFACE

After the earthquake of World War II and the spell of prostration which succeeded it, we are now facing the onset of an inevitable tidal wave of scientific publication. It will be some time before proper appraisal of this period is ossible, and all that can be done by the author of a text book is to select such xamples as represent, in his opinion, either current trends or well-established vances. For the ultimate benefit of reviewers, however, I shall not give in is preface a list of the main new features, but leave the matter to their dis-ernment.

Despite a good deal of pruning of older matter and compression of the new, some increase in size has been unavoidable: part of this is due to the inclusion of 60 new figures. These are:—Figs. 30, 31, 32, 51, 55, 62, 85, 89, 90, 101, 102, 103, 104, 114, 142, 172, 197, 199, 208, 213, 214, 241, 254, 265, 266, 295, 296, 297, 298, 299, 300, 301, 302, 303, 385, 395, 397, 398, 402, 406, 423, 427, 455, 456, 492, 499, 516, 534, 541, 543, 566, 567, 586, 586a, 595, 609, 618, 636, 637, 638.

I wish to convey my best thanks to authors and publishers, all of whom are duly acknowledged in the figure legends, for their permission to copy some of these figures; the remainder of the figures have been specially produced for is book. I have again to thank Professor H. Hartridge, F.R.S., for revision f the chapters on "Special Senses".

In accordance with the fashion in the various Services of Government, and erhaps as an omen of the shape of things to come, there has developed in late ears a considerable use of abbreviations in the form of initial letters. Although I do not like these, and have generally avoided their use in the text, I have bowed, or at least nodded, to the inevitable by giving a table of some of the common ones on p. xi.

I have to thank several readers for pointing out errors and misprints in previous editions and many workers in all parts of the world for valuable reprints which have helped to provide material or to shape my outlook.

My best thanks for reading galley proofs and for many suggestions are due to Dr. M. de Burgh Daly, Dr. Sheila Howarth, Dr. H. Ellis Lewis, Dr. O. C. J. Lippold and Dr. A. Schweitzer. To Dr. M. Grace Eggleton I once more express my gratitude for her reading of all page proofs, for many valuable suggestions and for preparation of the Index. I also wish to thank my Secretary, Miss Sybil Reid, for her competent and zealous work over many months in pre-aration of manuscript and in checking all proof corrections.

It has been a pleasure to work with the staff of Messrs. J. & A. Churchill, td., and I would especially record my gratitude to Mr. J. Rivers and to Ir. A. S. Knightley. The usual high standard of their artists and printers has een maintained.

C. LOVATT EVANS

Winterslow,
Wiltshire.
July 1st, 1949.

368.

CENTRAL FOOD TECHNOLOGICAL RESEARCH INSTITUTE, * MYSORE. *
LIBRARY

BOOK VII

The Removal of Waste Material and Temperature Conservation—Excretion and Temperature Regulation

CHAPTER PAGE
Historical Note 980
XLVII. RENAL EXCRETION 983
XLVIII. THE SKIN AND THE SKIN GLANDS 1027
XLIX. THE TEMPERATURE AND HEAT-BALANCE OF THE BODY 1034

BOOK VIII

The Provision of Special Chemical Stimulants, and of Measures for Continuance of the Species—The Endocrine Organs and Reproduction

Historical Note 1055

A. Endocrine Organs

L. THE ENDOCRINE ORGANS 1057

B. Reproduction

LI. THE FORMATION OF GERM CELLS 1095
LII. REPRODUCTION IN MAN AND OTHER MAMMALS . . 1103
LIII. THE SECRETION AND PROPERTIES OF MILK . . . 1144
LIV. HEREDITY 1152

SOME COMMON ABBREVIATIONS

A.C.H. = Adrenal cortical hormone.
A.C.T.H. = Adrenocorticotropic hormone.
A.D.P. = Adenosine diphosphate.
A.N.T.U. = Alpha-naphthylthiourea.
A.P.E. = Anterior pituitary extract.
A.T.P. = Adenosine triphosphate.
B.A.L. = British Anti-Lewisite (= 2 : 3 dimercaptopropanol).
B.M.R. = Basal metabolic rate.
B.P. = Blood pressure.
C.10 = Bistrimethylammonium decane diiodide, or decamethonium iodide.
C.R.O. = Cathode ray oscillograph.
D. = Diodrast.
D.B.T. = Dry bulb temperature.
D.C.A. or
D.O.C.A. = Desoxycorticosterone acetate.
D.O.P.A. = Dihydroxyphenylalanine.
D.P.N. = Dipyridino nucleotide.
D.T*m*. = Diodrast transport maximum.
E.C.G. = Electrocardiogram.
E.E.G. = Electroencephalogram.
E.T. = Effective temperature.
F.F. = Filtration fraction.
F.S.H. = Follicle-stimulating hormone.
G.F. = Glomerular filtrate.
I.U. = International unit.
I.C.S.H. = Interstitial cell-stimulating hormone.
L.H. = Luteinizing hormone.
M.Eq. = Milli-equivalent.
M.E.V. = Million electron volts.
M.L.D. = Minimal lethal dose.
M.T. = Methyl testosterone.
N.T.P. = Normal temperature and pressure.
P.A.H. = Para-amino hippuric acid.
P.G.A. = Pteroylglutamic acid.
P.F. = Plasma flow.
P.M.S. = Pregnant mare's serum (similar to F.S.H.)
P.U. = Pregnancy urine or chorionic gonadotropin.
P.R. = Pulse rate.
R.H. = Relative humidity.
R.Q. = Respiratory quotient.
S.D.A. = Specific dynamic action.
S.T.P. = Standard temperature and pressure.
T.Ac. = Testosterone acetate.
T.E.A. = Tetraethylammonium chloride.
T*m*. = Transport maximum.
T.P. = Testosterone propionate.
T.P.R. = Total peripheral resistance.
U.S. = Unconditional stimulus.
V.D.M. = Vaso depressor substance.
V.E.M. = Vaso excitor substance formed by kidney.
W.B.T. = Wet bulb temperature.

SOME UNITS AND CONSTANTS

Prefixes used in nomenclature, in addition to usual ones :—

Micro = 10^{-6} *i.e.*, one-millionth (μ).
Milli = 10^{-3} *i.e.*, one-thousandth, etc. (m.)
Centi = one-hundredth. (c.)
Kilo = 10^3 *i.e.*, one thousand. (k.)
Mega = 10^6 *i.e.*, one million, etc.
E.g., 1 Kilo calorie = 1 Calorie = 10^3 calories.
1 Micro calorie= 10^{-6} calories.

The following abbreviations and signs are employed for units :

mile	mi.	milligram	mg.
kilometre	km.	microgram (= 0·001 mg.)	μg. (or γ)
metre	m.	kilogram-metre	kg.m.
centimetre	cm.	gram-centimetre	g.cm.
millimetre	mm.	cubic metre (1,000 l.)	cu.m. (or
micron (10^{-3} mm.)	μ	litre	l.
millimicron (10^{-6} mm.)	mμ.	millilitre (= 1·000027 c.c.)	ml.
Angstrom unit (10^{-7} mm. or		cubic centimetre	c.c.
10^{-4} μ.)	A.U. or Å.	cubic millimetre	c.mm.
millionth micron (10^{-9} mm.)	μμ.	hour	hr.
square metre	sq.m. or m².	minute	min.
kilogram	kg.	second	sec.
gram	g.	millisecond (0.001 sec.)	msec. (or
centigram	cg.	microsecond (0·001 msec.)	μsec.

1 farad (F) = an electrical capacity which holds 1 coulomb when charged to 1 volt
microfarad = μF.

1 Microcurie (μC) = radioactivity of $3{\cdot}7 \times 10^4$ disintegrations per second.

1 Roentgen (*r*) = radiation producing 1 electrostatic unit of ions of both signs in
air at 0° C. and 760 mm. Hg.

1 cal. $\begin{cases} = \text{heat to raise 1 gram of water through } 1^\circ \text{ C.} \\ = 4{\cdot}2 \times 10^7 \text{ ergs} = 42{\cdot}7 \times 10^3 \text{ gram cms.} \end{cases}$

1 micro cal. = 42 ergs = $42{\cdot}7 \times 10^{-3}$ gram cms.

1 Cal (or 1 kilo cal.) = 1,000 cal. (= 4 B.Th.U. or 0·00004 Ther

Kinetic C.G.S. unit of force = 1 dyne = force which, acting on mass of 1 gram, ca
it to move with an acceleration of 1 cm. per second, per second.

Gravitational unit of force = 1 gram weight = 981·5 dynes.

Kinetic unit of work = 1 erg = 1 dyne-cm. = $23{\cdot}84 \times 10^{-9}$ calories = 10·16 ×
gram cms.

Gravitational unit of work = 1 gram cm. = 981·5 ergs = 23·4 micro cals. 1 kg.m
2·34 cal. (= 0·468 c.c. O_2 at 5 Cal/l.).

Kinetic Unit of Power = 1 erg. per second.

1 kilowatt = 10^3 watts = 10^{10} ergs per second.

Gravitational Unit of Power = 1 gram-cm. per second.

1 horse-power = 0·746 kw = 457 kg-m. per minute.

Mole (M) = 1 g. molecule = 6×10^{23} molecules. Molar = 1 g. molecule per l.

Millimole (mM) = M × 10^{-3} : Millimolar = 1 mg. molecule per l.

Micromole (μM) = M × 10^{-6}.

Osmotic pressure of molar solution (undissociated) = 22·4 atmospheres.

Osmotic pressure of binary electrolyte (whose degree of dissociation = δ) in
solution = $(1 + \delta) \times 22{\cdot}4$ atm. = $(1 + \delta) \times 22{\cdot}4 \times 760$ mm. Hg.

CHAPTER 1

INTRODUCTION

HE name "physiology," derived from a Greek root with the Latin equivalent *Physiologia*,* and originally connoting "natural knowledge," was first used n 1542 by Fernel,† but was not generally adopted until the 19th century. It now signifies the study of the phenomena presented by living organisms, the lassification of these phenomena, and the recognition of their sequence and elative significance, the allocation of every function to its appropriate organ, and the study of the conditions which determine each function.

Life is the indefinable of biological science, as time and space are indefinables or the physical sciences. By this is meant that we cannot even approximately define life without the idea itself being implicit in our definition. Our business s, given living things, to study their phenomena.

The fundamental phenomena of life are essentially identical in all living organisms. The interest of mankind was, however, naturally awakened in connection with his own body, and the science, growing up as ancillary to medical studies, has always taken man as its chief type of study. Further, the choice of the higher animals in general receives justification from the fact that, whereas morphology must proceed from the lowest to the highest organisation, the science of function presents its problems in their clearest form in the most highly differentiated organisms. In the unicellular animal all the essential functions of living beings are carried out, often simultaneously, in one ittle speck of protoplasm. It is only when, as in the higher animals, the chief parts of the body have highly specialised functions that it becomes possible to peer into the details of these functions with some chance of understanding them.

Organisation. What are the fundamental phenomena which distinguish iving things ? Their distinctive feature is perhaps *organisation*, and in the higher members this organisation becomes more and more distinct. This greater complexity of organisation runs parallel with increasing range and power of adaptation, attained by the setting apart of special structures (organs) or the performance of definite functions. The parallelism between the development of function and structure justifies us in the assumption generally, though often only tacitly, made by physiologists, that the structure is the determining actor for the function. Further, as living animals show progressively increasing differentiation of structure in all their parts, and with this an increased power f adaptation of each part, so by increasingly delicate co-operation between heir various parts, they display a progressive advance in *integration* which urther augments the powers of the individual to adapt itself to its surroundings.

Energy Expenditure. When dealing with the higher animals, we are nclined to lay stress on the phenomena involving a discharge of energy. The fe of a man in the ordinary sense of the term is made up of those movements hich place him in relationship with his environment. For these movements, s for the maintenance of a constant body-temperature, a continual expenditure f energy is necessary. Experience teaches us that these movements come to n end in the absence of food or of oxygen, and that an increased call on the nergies of the body must always be met by an increased demand for air and od. Two further processes must therefore be included among those making o our conception of life, viz. the function of assimilation (the taking in and digestion of food), and the function of respiration, by which oxygen is absorbed and carbon dioxide is excreted into the surrounding atmosphere.

---

* FULTON. *Yale J. Biol. Med.*, 1930, **3**, 59 ; *Science*, 1933, **78**, 109.

† SHERRINGTON, The Endeavour of Jean Fernel, 1946. (Cambridge Univ. Press.) *Arch. néerl. de Physiol.*, 1947, **28**, 369.

The substances which make up our foodstuffs are all capable of oxidation. In this process of oxidation there is liberation of heat. In the body a similar oxidation occurs, and energy is thus set free which is available for the activities of the living organism.

Before we can make any accurate investigations of the conditions which determine these activities, we must know whether the laws of the conservation of mass and the conservation of energy hold good for the processes within the living body. Thousands of experiments have been made, both on man and on animals, in which the total income of the body, viz. food and oxygen, has been weighed, and compared with the total output, viz. carbon dioxide, water, and urea, etc. In every case complete equality has been obtained.

In the determination of the equation in the case of the total energy of the body we have, in the first place, to measure the material income and output of the body, and to determine the total heat which would be evolved by the oxidation of the foodstuffs ingested to the carbon dioxide, water, etc., that are given out. We must then compare the figure so obtained with the *actual* output of heat by placing the animal in a specially constructed calorimeter, described in Chapter XLI. Rubner * thus measured in dogs the energy income and output of the body, and more elaborate experiments were made on men by Atwater and Benedict, a summary of whose results are given below :

INCOME AND OUTGO IN MAN DURING 143 DAYS
(Atwater and Benedict) †

Number of subjects used.	Condition.	Total number of days of experiment.	Average daily net income (calculated potential energy of material oxidized). Calories.	Average daily net outgo (energy given off by body). Calories.	Percentage difference.
4	Rest	41	2,246	2,246	0·0
3	Work	66	4,682	4,676	− 0·1
3	Special diet	26	2,290	2,305	+ 0·7
2	Special diet and work	10	3,719	3,702	− 0·5

The important deduction to be drawn from these observations is that the foodstuffs which are oxidised in the body develop in this process exactly the same amount of energy as when they are burnt outside the body.

From one aspect, therefore, the animal body may be looked upon as a machine for the transformation of the potential energy of the foodstuffs into kinetic energy, represented by the warmth and movements of the body as well as by other physical changes. In the living organism, however, it is not easy to distinguish between the source of energy and the machinery. When we endeavour to trace the foodstuffs after their entry into the body, we lose sight of them at the point where they apparently form an integral part of the living tissues. During activity there is a discharge of the products of oxidation of this living matter, which therefore becomes reduced in mass. This disintegration of the living matter, associated with activity, is always followed by a period of increased integration by the assimilation of more food. Our conception of life must therefore involve the idea of a constantly recurring cycle of processes.

---

* RUBNER. *Ztschr. f. Biol.*, 1894, **30,** 73. Gesetze des Energieverbrauchs bei der Ernährung. (Deuticke, 1902).

† ATWATER and BENEDICT. *U.S. Dept. Agricult. Bull., No.* **136,** 1903, p. 194.

one of *anabolism*, repair, or assimilation, and the other, associated with activity, of *katabolism* or dissimilation. Together they are called *metabolism*.

In the process of assimilation substances of high potential energy are formed, and this energy is obtained either from that imparted to the system at the moment of assimilation, as *e.g.* in the assimilation of carbon from carbon dioxide under the influence of the sun's rays, or from that contained in the foodstuffs themselves. In animal cells it is the latter method which is adopted and a foodstuff therefore connotes some substance which can be taken in by the cell and can serve it as a source of chemical energy. In assimilation there may also be a synthesis of more complex from less complex compounds, without their necessary entry into the structure of the living molecule. Assimilation requires the ingestion of food into the organism, and then its digestion, *i.e.* its solution in the juices of the cells. These two processes are often succeeded by an actual growth in the living material. Digestion is effected in most cases by the production of solutions containing *enzymes*, *i.e.* agents which have the power of hydrolysing the different foodstuffs.

In the process of dissimilation in most organisms the energy for their activities is derived from oxidation. A necessary condition therefore is the presence of oxygen. As a result of this oxidation, products are formed which are of no further value and are therefore *excreted*, *i.e.* turned out of the cell. The chief of these are carbon dioxide and water. There are also many substances resulting from the oxidation of nitrogenous substances which have to be excreted.

Although the need for oxygen is so general a quality of living protoplasm, its presence does not seem to be necessary for all kinds of life. Indeed, most cells can survive for a short time in its absence, and can even carry out oxidative processes by indirect reactions. Further, many bacteria are known which are anaerobic, *i.e.* exist only in the absence of oxygen. Examples of such are bacillus tetanus and the bacillus of malignant œdema.

The activities of a living cell or organism can be regarded in every case as dependent originally on environmental change, and are of such a nature that they tend to preserve the organism intact. The property of reacting in such a manner to changes in the environment is fundamental to all protoplasm and is spoken of as *excitability*, and any change which will influence an organism and cause a response in it is known as a *stimulus*. Stimuli may be of various kinds. Thus mechanical, thermal, chemical, electrical changes, light and so on may act as stimuli. The reactions which they evoke in every case involve changes in the metabolism of the cell. Sometimes this change may be assimilatory in character, or at any rate a reduction of dissimilation. In such a case the stimulus is spoken of as inhibitory, because it diminishes or prevents the output of energy by the organism. The frequent result of a stimulus is an increased output of energy, which may appear in the form of movement, heat, or chemical change.

A common feature of all dissimilatory changes evoked by the application of a stimulus is that the energy of the response is many times greater than the energy represented by the stimulus, the excess, of course, being supplied at the expense of the potential energy of the material of the living protoplasm.

In unicellular organisms and single cells, definite classes of motor reaction to stimulus have been described. Among the primitive reactions of cells perhaps the most important in the life of higher animals are those grouped under the term *chemiotaxis*. Thus, the movements of the white corpuscles of the blood are largely determined by their chemical sensibility, and various substances can be divided into (*a*) those which exercise positive and (*b*) those which exercise negative chemiotactic influence on the leucocytes. The introduction under the skin of an animal of a capillary tube containing a solution of substances of the first class, such as peptone, "leukotaxin," or the chemical products of certain bacteria, leads to an accumulation within the tube of leucocytes which pass to it from all the surrounding tissues. Other substances, such as quinine, exert a negative chemiotaxis. Tubes filled with these, after introduction into the subcutaneous tissue of a mammal, will be found many hours later to contain no leucocytes at all.

Growth, Reproduction. If the anabolic process predominates, we obtain growth. Closely associated with growth is the power possessed by all living organisms of repair, *i.e.* the replacement by newly formed healthy material c parts which have been damaged.

The process of growth does not, in the individual, proceed indefinitel At a certain stage in its life every organism divides, and part of its substanc is thrown off to form one or more new individuals, each of them endowed wit the same properties as the parent organism. In all the higher forms the paren organism begins to undergo decay, until finally death takes place.

Adaptation. All these phenomena, viz. organisation, assimilation respiration, the discharge of energy, growth, reproduction, and death itself are bound up in our conception of life. All living organisms have one featur in common, viz. they are endowed with the power of *adaptation.* Adaptatio is " the continuous adjustment of internal relations to external relations." living organism is a highly unstable system which tends to undergo disintegration if its average environment is changed beyond certain narrow limits. The real environment of the living cells is the tissue fluid in which they live ; this *internal environment* tends to be altered by changes in the general environment in which the animal is placed. It is evident that the condition for survival of the organism must be that changes in the environment are counteracted so that, whatever the general environment may be, the internal environment stays within viable limits. This is what is meant by adaptation.

We have no knowledge of how living things first came into existence, but having once arisen in a primitive form, the " adaptation," *i.e.* the reactions of the primitive living material to changes in its environment, must become ever more and more complex, since only by means of increasing variety of reaction is it possible to provide for the stability of the system within greater and greater range of external conditions. The difference between higher and lower forms is therefore one of complexity of reaction, or of range of adaptation.

Organism and Environment. Though it is a philosophical point on which there is no need to dwell, it should be emphasised that organism and environment must be viewed as a whole. But so far as possible the organism adjusts itself to any changes in its environment or " adjusts its internal to its external relations."

This " adjustment of internal to external relations " is possible only within strictly defined limits, limits which increase in extent with rise in the type of organism, and in the complexity of its powers of reaction. Among the chief of them are temperature, and the presence of food material, water and oxygen. Many organisms are killed by the alteration of only a few degrees in the temperature of their environment. In the higher animals a greater stability in face of such changes has been accomplished by the development of a heat-regulating mechanism, so that, provided sufficient food is available, the temperature of the body is maintained at a constant level. In the primitive condition, the food material must be of a given character and form a constant constituent of the surrounding medium. In the higher forms, the development of a complex digestive system has enabled the organism to utilise many different kinds of food, while the storage of excess as reserve material provides for a constant supply of food to the cells of the body, even when it is temporarily wanting in the environment. Moreover, but few cells of the higher organisms are called upon to face the vicissitudes of light, temperature, chemical composition, etc., of the general environment : instead, most of the cells of the body live in darkness and in contact only with that inner environment, the blood and tissue fluid, the composition and temperature of which are kept constant by wonderfully adapted arrangements. The name *homeostasis* (homoiostasis) is given by Cannon to the condition of uniformity which results from the adjustments of living things to changes in their environment.

To sum up. Our objects in the study of physiology include the description of the chief reactions of the body to changes in its environment, the analysis of these reactions into the simpler reactions of which they are made up, and the assignment to each differentiated structure of the organism its part in every reaction. We must determine the conditions under which each response takes place, so that we may learn to evoke it by application of the appropriate change of environment.

In attacking this problem our methods cannot differ fundamentally from those of the physicist and chemist. In every case our experiments will consist in the observation and measurement of processes of one kind or another which we shall hope to interpret in terms of mass or energy. Physiology, if it could be completed, would therefore describe the *how* of every process in the body. It would state the sequence of events and would summarise these as so-called "laws." These laws would, however, no more explain the phenomena of life than does the "law of gravitation" explain the fact that two masses attract each other with forces directly proportional to the product of their masses. Nor can we hope to explain physiological phenomena by reference to the laws of physics and chemistry, since these themselves are only expressions of sequences, and not explanations. With every growth in science, however, its generalisations become wider and its laws summarise ever more extensive groups of phenomena. The principle of adaptation is the only formula which will include all the phenomena of living beings. This principle must provisionally be accepted as fundamental by the biologist, as the physicist accepts the first law of thermodynamics.

The consciousness of the inadequacy of our explanations, which must be experienced with great force the more deeply the physiologist endeavours to peer into the processes within the living cells, at one time led some to the assumption of a special quality in living organisms which was designated as "vital activity." Such views are classified together under the term "*vitalism.*" As a working hypothesis it must be sterile because there is no evidence that the introduction of vitality enables the known laws of nature to be overruled in any way: whatever the living thing accomplishes, it may do so with the aid of quite orthodox phenomena. That the presence of life does, however, result in an orderly marshalling of these orthodox phenomena, and that in a quite peculiar way, could hardly be questioned. This again, is a philosophical point and need not detain us further than to note the apt remark of Claude Bernard, that "vital force directs phenomena which it does not produce; physical agencies produce (in living things) phenomena which they do not direct." In many cases, however, the terms "*vitalism*" and its antithesis "*mechanism*" are used unjustifiably. The phenomena of living matter are not arbitrary, and since the conditions are infinitely complicated, they are incapable of arbitrary simplification when compared with those concerned with the study of non-living matter, and are hence more difficult to analyse or to control. The physiologist, basing his belief in the principle of adaptation must reject vitalistic "explanations" and seek to discover by experiment the conditions which determine the appearance of phenomena in living things. But he must equally reject an arrogant and superficial doctrine of "mechanism" which would suppose that life can be the outcome of any chance encounter of physical and chemical phenomena, and here again the principle of adaptation must be his guide in refusing to accept the suggestion that physiology is nothing more than applied chemistry and physics. Holding these aims always before him, he must abide by the answer he receives from his experiments as to the ways and means of adaptation, and can hope for little more.

Throughout this chapter we have admitted no necessary dividing line between the different classes of phenomena in the conceptual universe, although in the present state of our knowledge we are far from being able to include the whole of them under the same general laws. It might be objected that in taking up this attitude we have left out of account one supreme fact, viz. the existence of *consciousness* in ourselves. As a comparative and objective study, however, physiology is concerned, not with the study of consciousness, but with the conceptions in consciousness of the *phenomena* presented by living beings.

THE METHOD OF PHYSIOLOGY. Physiology derives most of its inspiration from observations incidental to the phenomena of our daily life and from those made in the course of medical practice. In the prosecution of his enquiries the physiologist often must make appeal to other sciences, but especially to chemistry and physics. In the first place the chemistry and physics of the various separate organs of the animal body must be investigated. This involves first an accurate knowledge of the gross and minute structures, and hence requires the aid of anatomy and histology. Here very frequently, too, comparative anatomy and histology furnish valuable clues to function.

Then the chemical, physical and physico-chemical build must be taken into consideration with a view to the determination of the changes which these exhibit under altered conditions of environment, chemical or physical in nature.

In this way analytical physiology is built up, by experiments, carried out on living animals, or on separated parts of them surviving for a time under suitable conditions ; knowledge of the properties and functions of the various structura parts is obtained. Then, in synthetic physiology, we attempt to find how, by integrative processes, the functions of the body as a whole are synthesised from the co-operative activity of the various parts, and how the organism as a whole reacts, by appropriate adjustment of its individual organs, to changes in its external environment.*

"*Mechanisms*" *in Physiology.* In his physiological studies the student will often meet with the word "mechanism," *e.g.* "the mechanism of gastric secretion"; ". . . of walking"; ". . . of reflex action," etc. This rather overworked word need not be taken too seriously, because as a rule it has no more significance than the equally busy and now notorious adjective "marked." "Mechanism of" really means, as a rule, "the phenomena concerned in . . ." as *e.g.* "the mechanism of bone formation," though in other instances, which the reader can readily find as *e.g.* "The mechanism of speech," the word is more literally employed.

The Normal. Individuals of the same species resemble one another more or less closely in structure and in their physiological properties, and this tendency of biological characters towards a standard type is an expression of the approximate harmony between the organism and its environment. The resemblance is greatest between animals born of the same parents at the same time. As species further apart are considered the differences become greater, and, between the extremes of the animal kingdom, are so great that comparisons become vague.

Yet, even when the relationships between individual animals are of the closest kind, *e.g.* in the human race, these do show differences from one another. It is a duty of human physiology to study the normal human individual, since this provides an important baseline for the study of disease ; yet it is far from easy to state what is the normal of any given character. Further, in all physiological experiments the results must show a good deal of apparently arbitrary variation depending on the individual animals used. Hence the importance of proper controls, and the decision of what constitutes a significant result, cannot be overestimated. Often, in order to be certain that representative results are being considered, it is necessary to repeat experiments many times. The error of observation which is usually slight in the exact sciences, is often great in the biological ones, on account of the variability introduced by the use of varying individuals. The results are, however, capable of being treated by the application of the same laws of probability as are employed in other sciences, *i.e.* they can be treated statistically. Caution is especially necessary in drawing conclusions from a limited number of experiments, as to whether the results obtained differ from those of a control series or not. Thus if a given result happens 4 times out of 8 in one series and 7 times out of 16 in another, we are not justified in concluding that it happens more often under the conditions of the first, because chance could easily account for the difference. But 4 out of 8 in one series and 3 out of 16 in the second would be a significant difference.

It is a common practice to speak of averages, that is, of arithmetic means, as one way of overcoming the difficulty, *e.g.* we can say that the average number of red cells per cubic millimetre of blood in a given community of men is 5·2 millions. Such a method, though the simplest and the one most commonly used, is often unsatisfactory, because it does not tell us how great or how relatively frequent are the departures from the mean, so that we should have no means of knowing in the example quoted whether 4·5 millions was within the normal range or not. There is consequently a growing practice of using more accurate means for the expression of statistical results.

One of the methods employed is the use of the *frequency diagram*, an example of which is given in Fig. 1, which illustrates data regarding the stature of a large number of individuals. These have been divided into frequency groups differing in stature by 1 inch. The height of each rectangle is proportional to the number of individuals having the height shown at the abscissa, the grouped figure being called a histogram. A smoothed curve, or frequency curve, drawn from this, shows the frequency with which any stature between the given limits occurs.

---

* As an example of essays in synthetic physiology may be quoted J. Barcroft's "Features in the Architecture of Physiological Function," Cambridge, 1934.

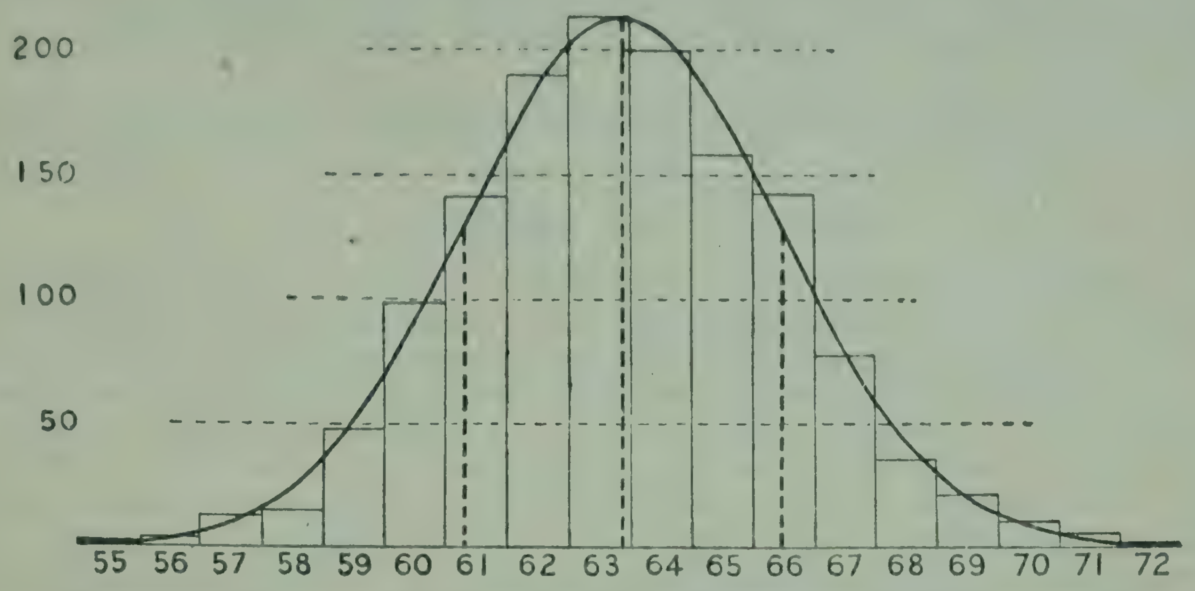


FIG. 1. Frequency diagram of distribution of stature of 1,375 women, in groups differing by one inch. (Pearson and Lee.) (From Fisher's "Statistical Methods for Research Workers." London. Oliver & Boyd.)

(Other examples are given in Chapter XXIX showing distributions for blood counts, and in Chapter XXXIII showing blood pressures.)

The *standard deviation* or mean error, is one of the means of improving the value of results; it measures the amount of variation in the individual figures from which the average was obtained. The accuracy of an average result depends on the square root of the number of observations made, so that, in a series of n determinations with individual differences summed up as Σd, it can be shown that the standard deviation of a single observation is $\pm \sqrt{\frac{\Sigma d^2}{n-1}}$, while the standard deviation of the average is given by the formula $\pm \sqrt{\frac{\Sigma d^2}{n(n-1)}}$.

To take an illustration given by Burn, we may consider the investigation of the lethal dose of a drug, as determined on a series of five cats, which gave the result as 17·44 ± 1·67 c.c. The figure 1·67 is the standard deviation of the average, and means that, "if a number of determinations of the average lethal dose be made, each time on five cats, then twice out of three times the difference between the value obtained and the true value will be less than 1·67, and once out of three times this difference will be greater than 1·67." If we take twice the standard deviation, and call the result 17·44 ± 3·34 c.c. the report will probably be wrong only once in twenty-two times; or if three times the S.D. or 17·44 ± 5·01 c.c., then it will probably only be wrong once in 370 times.

A difference between two series of similar experiments with standard deviation $= \sigma$ may be considered to be significant if, the two series having given results of $x_1 \pm \sigma_1$, and $x_2 \pm \sigma_2$, the ratio $\frac{x_1 - x_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$ is greater than 3.

The *probable error* of a mean is given by the formula $\pm\ 0{\cdot}6745\ \frac{\sqrt{\Sigma d^2}}{n\,(n-1)}$ or, by the more convenient approximation of Peters, the probable error of the arithmetic means of a series of observations $= \pm \frac{0{\cdot}8455\ \Sigma(+\,d)}{n\sqrt{(n-1)}}$ where $\Sigma\,(+\,d)$ is the sum of the deviations of every observation from the mean, their sign being disregarded. The probable error determines the degree of confidence we may have in using the mean as the best representative value of a series.*

For Reference

ADOLPH (1943). Physiological Regulations. (Lancaster, Pa.: Jaques Cattell Press.)
BERNARD, CLAUDE (1855) (Translation 1927). An Introduction to the Study of Experimental Medicine. (Macmillan.)

---

* For further information on this specialised subject, see the book by FISHER and the articles by BURN and DUNN. A brief account is given in LAMB's "Introduction to Human Experimental Physiology." (Longmans, Green, 1930.)

BERNARD, CLAUDE (1878). Leçons sur les phenomènes de la vie. (Paris; Baillière.)
BURN, J. H. (1930). The Errors of Biological Assay. *Physiol. Rev.*, **10,** 146.
CANNON, W. B. (1929). Organization for Physiological Homeostasis. *Physiol. Review* **9,** 399.
DUNN, HALBERT L. (1929). Application of Statistical Methods in Physiology. *Physiol Rev.*, **9,** 275.
FISHER (1932). Statistical Methods for Research Workers. (London : Oliver & Boyd.)
HALDANE, J. S. (1917). Organism and Environment. (Yale Univ. Press.)
LOEWENTHAL & WILSON. Problems of Chance in Medicine and Research. *B.M.J.*, 1939 (ii.) 110.
SCOTT, E. L. (1927). What constitutes an adequate series of Physiological Observations ? *J. Biol. Chem.*, **73,** 81.

HISTORY OF PHYSIOLOGY

As it is impossible to realise the present position of any science without some knowledge of the history of its development, the serious student of physiology should acquaint himself with the outlines of its history. By way of introduction, some very brief historical notes, dealing with the development of various branches of physiology, are given here and there in this volume. As regards the growth of the subject as a whole, it must be remembered that no part of scientific knowledge can develop alone, but that all parts depend on each other. For instance, until a reasonable body of anatomical knowledge had been established, no physiology, as we understand it, was possible; and until physics and chemistry had reached a certain state of development, much of the early physiology could be but tales of mystery and imagination.

Starting, then, with the foundation of modern anatomy by the publication, in 1543, of the *Fabrica Humani Corporis* of Andreas Vesalius, we may put down the following names of key men whose work enabled expansion of our subject to take place in one direction or another. They form important landmarks.

Name.	Life.	Important publication.	
		Date.	Significance.
Vesalius	1514–64	1543	Start of modern anatomy.
Galileo	1564–1642	c. 1600	Start of modern physics.
Harvey	1578–1667	1628	Experimental method in biology.
Boyle	1627–91	1660	Experimental method in chemistry.
Malpighi	1628–94	1661	Microscope used in biology.
Mayow	1645–79	1668	Application of chemistry.
Newton	1642–1727	1687	Development of modern physics.
Boerhaave	1668–1738	1708	Application of chemistry and physics.
Haller	1708–77	c. 1760	Important text-book on physiology.
Lavoisier	1743–94	1775	Combustion and respiration related.
Joh. Müller	1801–58	c. 1834	Important text-book.
Schwann	1810–82	1839	Cell theory established.
Bernard	1813–78	1840–70	Great experimentalist.
Ludwig	1816–95	1850–90	Great experimentalist. Introduced graphic method.
Helmholtz	1821–94	1850–90	Applications of physics.

The first journal devoted to physiology, first appeared in 1795; the English *Journal of Physiology* in 1878, and the *American Journal of Physiology* in 1898. The first English chair of physiology, as distinct from anatomy, was founded at University College, London, in 1874, and the first in the United States of America, at Harvard, in 1876.

It is thus seen that the subject of Physiology, in anything like the form we now know it, began to take shape little more than a century ago, and so is a young subject. Biochemistry is still younger, and has mainly developed in the lifetime of many not yet old, as an outgrowth of physiology.

For more details of the history of physiology, the student may consult the following works :—

FOSTER (1901). Lectures on the History of Physiology. (Cambr. Univ. Press.)
FRANKLIN (1933). A Short History of Physiology. (John Bale.)
FULTON (1930). Selected Readings in the History of Physiology. (Springfield : Thomas.)
FULTON (1931). Physiology. (Clio. Medica Series, Hoeber.)
SINGER (1928). A Short History of Medicine. (Oxford.)

BOOK I

GENERAL PRINCIPLES

A. Structural Principles. B. Biochemical Principles.
C. Biophysical Principles.

A. STRUCTURAL PRINCIPLES

CHAPTER II

THE STRUCTURAL BASIS OF THE BODY*

THE CELL

A cell is a small mass of protoplasm containing a nucleus. Since the formulation of the *cell theory* of life by Schwann and Schleiden about 1840, it has been customary to regard the function of any organ, or indeed of the whole body, as the sum of the functions of its constituent cells. The cell is therefore a physiological as well as a structural unit.

Many of the lower forms consist only of a single cell. In this cell are represented all the adapted reactions which we associate with life.

The general characteristics of cells present important similarities, whether they form the whole of an organism or but an infinitesimal part of a highly developed animal.

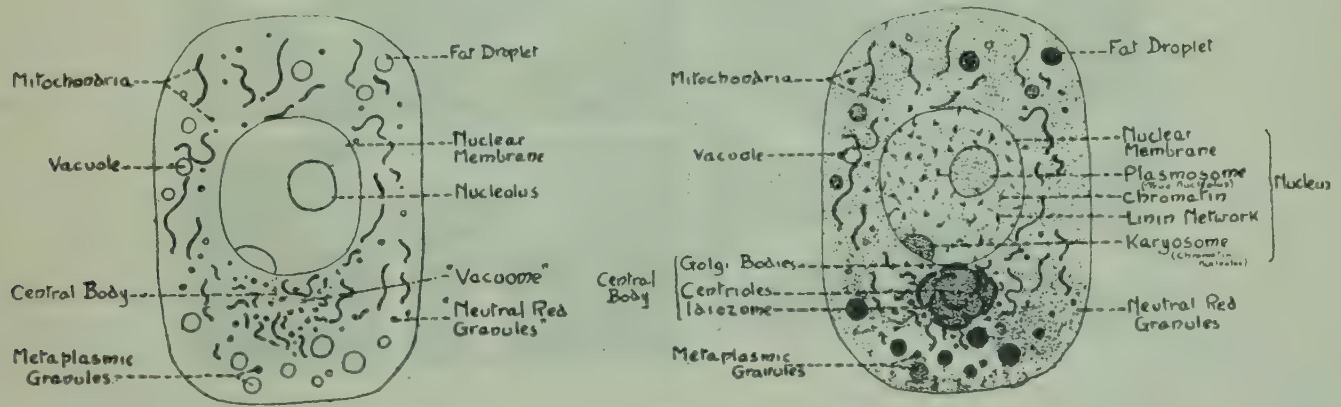

Mitochondria
Vacuole
Central Body
Metaplasmic Granules
Fat Droplet
Nuclear Membrane
Nucleolus
"Vacuome"
"Neutral Red Granules"
Mitochondria
Vacuole
Central Body
Golgi Bodies
Centrioles
Idiozome
Metaplasmic Granules
Fat Droplet
Nuclear Membrane
Plasmosome
(True Nucleolus)
Chromatin
Linin Network
Karyosome
(Chromatin Nucleolus)
Nucleus
Neutral Red Granules

Fig. 2. Diagram to show contents of typical cell. On left, structures which can be seen under appropriate conditions in the living cell. On right, the structures demonstrable by various fixing and staining methods. (Ludford.)

In the multicellular animals there are cells of many different kinds of structure and function. In this way there is a division of labour. The organism, biologically, is something more than the mere sum of its component cells, since it directs the growth and functions of these to its own ends, just as a nation is, biologically, something more than a large number of individuals. We must regard the organism also as a biological unit of definite significance.

Every cell, so far as we have been able to ascertain, is derived from some pre-existing cell, by a process of division, and every nucleus similarly from a parent nucleus. These important laws, so significant in the study of heredity, were expressed by Virchow (1855) and Flemming (1882) in the dicta "*omnis*

---

* The matter of this chapter will be familiar to students who have studied Histology, Cytology, and Embryology, and only gives such broad outlines as to recall the importance of those subjects to Physiology.

cellula e cellula" and "*omnis nucleus e nucleo*" respectively. As a necessary consequence it follows that the cytoplasm of the cell is also derived from the previously existing cytoplasm of the cell from which it arose.

The term "protoplasm" generally implies any material which we can regard as endowed with life. This material is very complex and of very varied composition. The organisation and reactions of every cell are the resultant of the molecular structure of the matter of which it is built up. The limited methods of the chemist show him that the composition of the muscle cell is entirely different from that of a leucocyte. The finer methods of the physiologist show him that every sort of cell has its own peculiarities of reaction to changes in its surroundings, and the reactions of the whole organism are but the sum of the reactions of its constituent cells.

Even in a single cell there is more than one kind of protoplasm, since the nucleus differs widely from the cytoplasm. Both are necessary for the life of the cell and both must be considered as "living." In the cytoplasm itself we find structures or substances which we must regard as on their way to protoplasm or as products of the breakdown of protoplasm (Fig. 2); but in many cases it is impossible to say whether a given material is to be regarded as living. Even in a single cell we may have differentiation among its different parts, one part serving for the process of digestion while others are employed for the purpose of locomotion, etc.

METHODS EMPLOYED IN THE STUDY OF MINUTE STRUCTURE

The relations between structure and function are so intimate that physiology would be incomplete if it did not take considerations of structure into account. Moreover, there has grown up recently an important branch of physiological investigation, namely, the enquiry into the physiology of individual cells, in which methods of histological research are in constant use, and are supplemented by a variety of instrumental techniques.

Routine histological technique has its limitations. We are often in doubt as to how far the structures we see in fixed and stained preparations have any real existence in the living cells. A method of fixation which produces minimal change of structure and physicochemical arrangement is to dry the tissue while in a frozen state.* A further difficulty is that the power of the microscope to resolve details has been developed to the uttermost limit, namely, to enable us to see structures whose dimensions are 0·25 μ, equal to half the wave-length of light. Some further resolution can be got by using light of shorter wave-lengths; more detail will be seen with green illumination than with yellow light, and by photomicrography with invisible ultra-violet light details down to 0·1μ can be made out. A further result of using ultra-violet light has been to show that certain structures, *e.g.* chromosomes, are more opaque to it than others, though both types may be equally transparent to the visible rays; hence not only finer structures, but actual new structures, may be revealed.

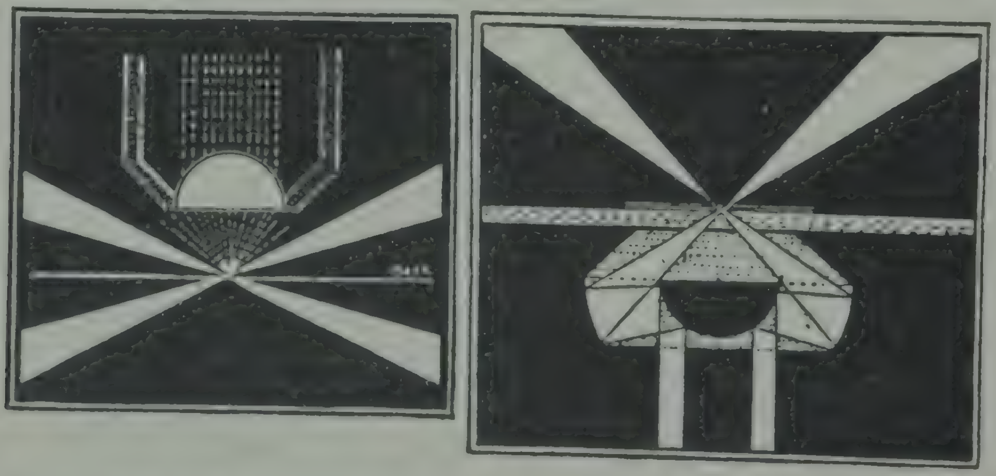


FIG. 3. To illustrate the principle of dark-ground illumination, and the method by which this is obtained with the use of a special substage condenser. (Courtesy of Messrs. R. and J. BECK.)

A further refinement is by the use of the principle of the ultra-microscope in what is called *dark-ground illumination*. With the aid of a paraboloid condenser the light is thrown nearly transversely across the preparation (Fig. 3), so that the field is dark except where particles lying in the path of the beam are able to reflect the light into the microscope (Fig. 3). This method will reveal the position of particles in living protoplasm, and give us some idea of their size, even though they are so small as to be invisible by any direct microscopic examination. Yet another useful technique is provided by a special method of illumination called *phase contrast* (Zernike), which develops contrast due to differences of refractive index, not normally detectable in the structure of living cells.* The reflecting microscope, in which lenses are replaced by curved mirrors, offers other advantages, *e.g.* absence of chromatic aberration, availability for any wave-length, including the ultra-violet, and long working distance, which facilitates manipulation under high powers.†

---

* BENSLEY *et al. Anat. Rec.* 1933, **57**, 205, 217, 369; **58**, 1; 1934, **60**, 251, 449.

Accelerated cinemicrophotography is a powerful aid to the observation of the changes taking place in living cells. Its use reveals the fact that many cells exhibit extensive changes in shape and position, inappreciable to unaided observation.

Much valuable information has been obtained with regard to the structure and behaviour of living cells by means of *micro-manipulation*, which has been extensively used as a means of dissecting the live cell or of injecting various fluids into it. The living cells are suspended, as a hanging-drop preparation, from a cover-glass, and are manipulated and dissected by means of two fine glass needles, the movements of which in all directions are controlled by micrometer adjustments (Fig. 4): injection is performed by fixating the cell with a needle and then perforating it at the desired place with a fine glass capillary through which the fluid is then forced from a micrometer syringe.

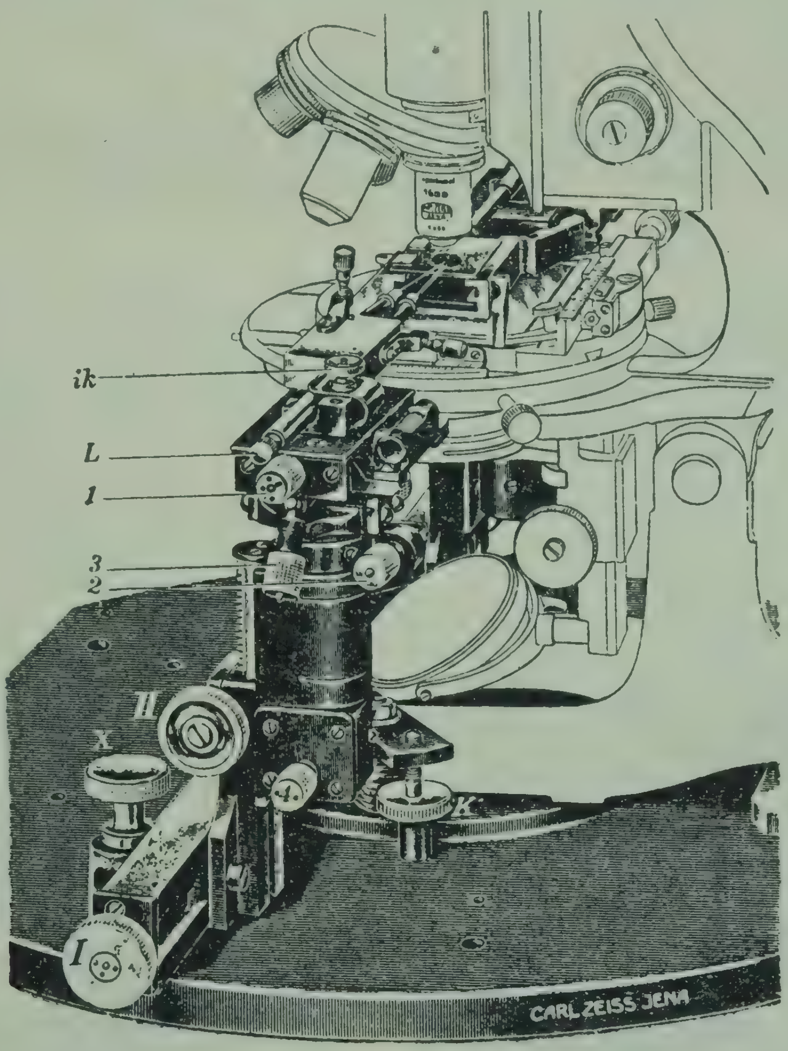


Fig. 4. Micro-manipulator in position. The manipulations are effected by six screws, two of which transmit coarse motions to the operating stands as a whole, while the other four actuate the micro-instruments. In the figure a double needle is shown on one stand and a single one on the other. (Courtesy of Messrs. Carl Zeiss.)

Another method which has found many applications of late is that of *tissue culture*, in which small portions of embryonic tissue are taken and kept at suitable temperatures in a nutrient fluid, such as a mixture of embryo extract, blood plasma and saline solution. The whole preparation must be aseptic, and is usually made as a hanging drop. The tissue continues to live, and under certain conditions the cells grow and multiply in an apparently normal manner until the cell mass reaches a certain size, when growth ceases and degeneration sets in unless a small piece of the tissue is taken and used to start another similar culture. In some cases the process of sub-culture can apparently be repeated indefinitely. Much has been learned of the structure, growth, and differentiation of embryonic cells by this method.

The histological staining methods are very usefully supplemented by a variety of so-called *intra-vitam staining* methods in which the dye is presented to the living organism, and *supra-vital staining*, in which the dye is added to a solution in which surviving tissues or isolated cells are suspended. Certain portions of cells have an affinity for particular dyes. Thus the mitochondria are stained in the live cell by Janus green, nerve material by methylene blue, many cell granules by neutral red, certain connective tissue cells by Pyrrol or Trypan blue, etc. This gives a valuable method for classifying cells, as well as for obtaining information with regard to their structure.

---

* Barer. *Q.J. Micr. Sci.*, 1947, **88**, 491.

† Barer. *Brit. Science News*, 1948, I, 12.

A good deal can be learned about the distribution of chemical substances within cells by means of various *microchemical reactions*. Thus the presence and location of fat, protein, glycogen, potassium, iron, etc., can be discovered by appropriate colour reactions applied to sections of tissues or to isolated cells. The distribution of the inorganic constituents can be studied by means of *micro-incineration* ; * sections of the tissue are incinerated on the slide, and on dark-ground illumination the deposit of ash is seen to vary in density in different kinds of cell or different parts of each cell.

Ultrastructure of Protoplasm. In its last resort the study of structure must become that of molecular orientation, and hence lies beyond the reach of ordinary microscopes. The electron microscope,† giving a resolution down to 4 mμ (·004μ) or less, has already found some applications, and the optical properties to polarised light have also been used. The latter shows whether molecules or micellæ are orientated in regular patterns. By the use of X-ray diffraction methods also, information regarding the molecular components of tissues can be obtained.‡ This gives resolution from 50 mμ down to 0·01 mμ. It therefore enables molecules to be studied. Some disease viruses have molecules 150 mμ long ; some proteins 5 mμ ; a water molecule 0·3 mμ. The methods are complicated and the interpretation of results, though promising, is still largely tentative.

For Reference

SCHMITT (1939). The Ultrastructure of Protoplasmic Constituents. *Physiol. Rev.*, **19**, 270.
DEMPSEY and WISLOCKI (1946). Histochemical Contributions to Physiology. *Physiol. Rev.*, **26**, 1.

STRUCTURE OF THE CELL. Cells are often enclosed by one or more outer coats not essential to life ; but there is always one indispensable layer, called the *plasma membrane*, which lies immediately outside the cytoplasm and represents its real boundary ; within this again is a thin shell, the *ectoplasm*, generally in gel form, and homogeneous. Inside this is the cytoplasm proper, within which is the nucleus.

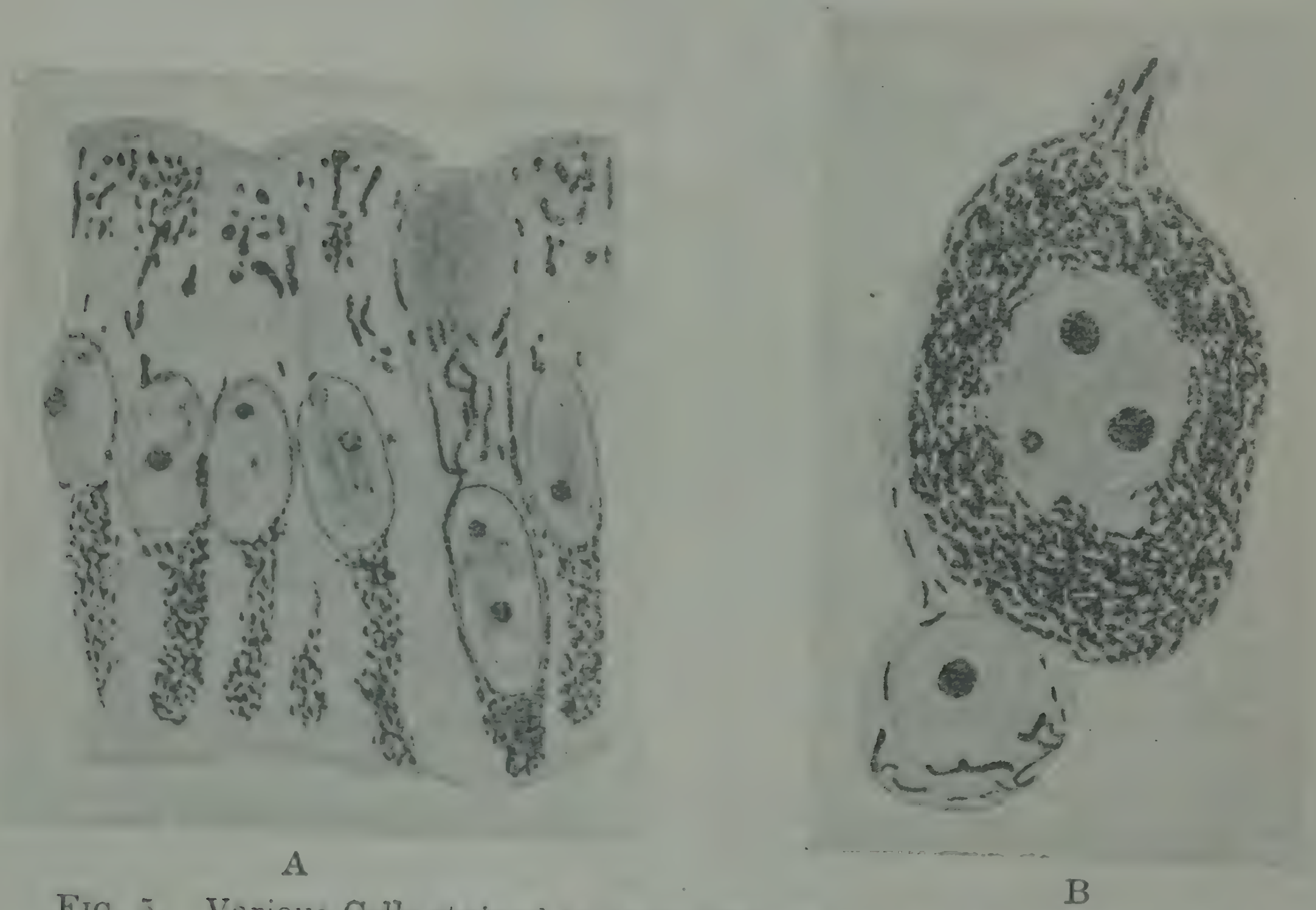


FIG. 5. Various Cells stained to show Mitochondria. (E. V. COWDRY.)
A. Lining Cells of the Small Intestine.
B. Two Nerve Cells, with different forms of Mitochondria.

The nucleus is generally bounded by a definite *nuclear membrane*. The contents of the nucleus are probably fluid in life, but on damage, death and coagulation rapidly occur. Within the nucleus are one or more smaller bodies called *nucleoli* ; in the living non-dividing nucleus no other structures are usually visible. After fixation various artefacts appear in the nucleus, *e.g.* masses or filaments of a material known as *chromatin*, which are strung, so to speak, on a fine network of material known as *linin*. The material filling up the meshes of the network is the nuclear sap or *nucleoplasm*. When the cell is about to divide, the chromatin becomes aggregated into filaments called *chromosomes*, and these are visible in the living cell.

---

* POLICARD. *Protoplasma*, 1929, **7**, 464. SCOTT. *Anat. Rec.*, 1931, **48**, **31**.
† ZWORYKIN *et al.* Medical Physics. Ed. Glasser, 1944, p. 387.
‡ CLARK. Medical Physics. Ed. Glasser, 1944, p. 1335.

The cytoplasm, which varies greatly in extent in different cells, varies also in its appearance, especially after fixation, being sometimes homogeneous, sometimes alveolar, sometimes granular in structure. In it can often be distinguished differentiated parts which may be regarded as organs of the cell.

The most universal of these cell inclusions, or possibly cell organs, are the *mitochondria*. These occur in the form of granules, rods or filaments in almost every type of cell, animal and vegetable (*cp.* Fig. 5). Their function is unknown, though they seem to be more numerous and longer in cells which are very active. Their presence must largely increase the surface of phase-boundaries within the cell, and it has been suggested that important chemical reactions occur at these surfaces of contact between mitochondria and cytoplasm. The important cytochrome oxidising systems are also located in the mitochondria. They are often grouped at one pole of the cell; thus in gland cells they are frequently found at the pole nearest to the blood capillaries and farthest from the lumen. They are more abundant in young than in old cells, and in the living cell are seen to move about readily in the cytoplasm. When the cell divides the mitochondria are shared between the two daughter cells. In some conditions the mitochondria can be dragged out of the cell by microdissection methods, hence they are relatively solid. They contain some 40 per cent. of fatty substance, and two different proteins, but are said to contain no lecithin or cephalin.* In the living cell they can be stained with Janus green. Usually they are dissolved by alcohol and by weak acids, and so are not displayed in cells hardened to display nuclear structure. They are best preserved by hardening in formalin and bichromate and staining with iron hæmatoxylin.

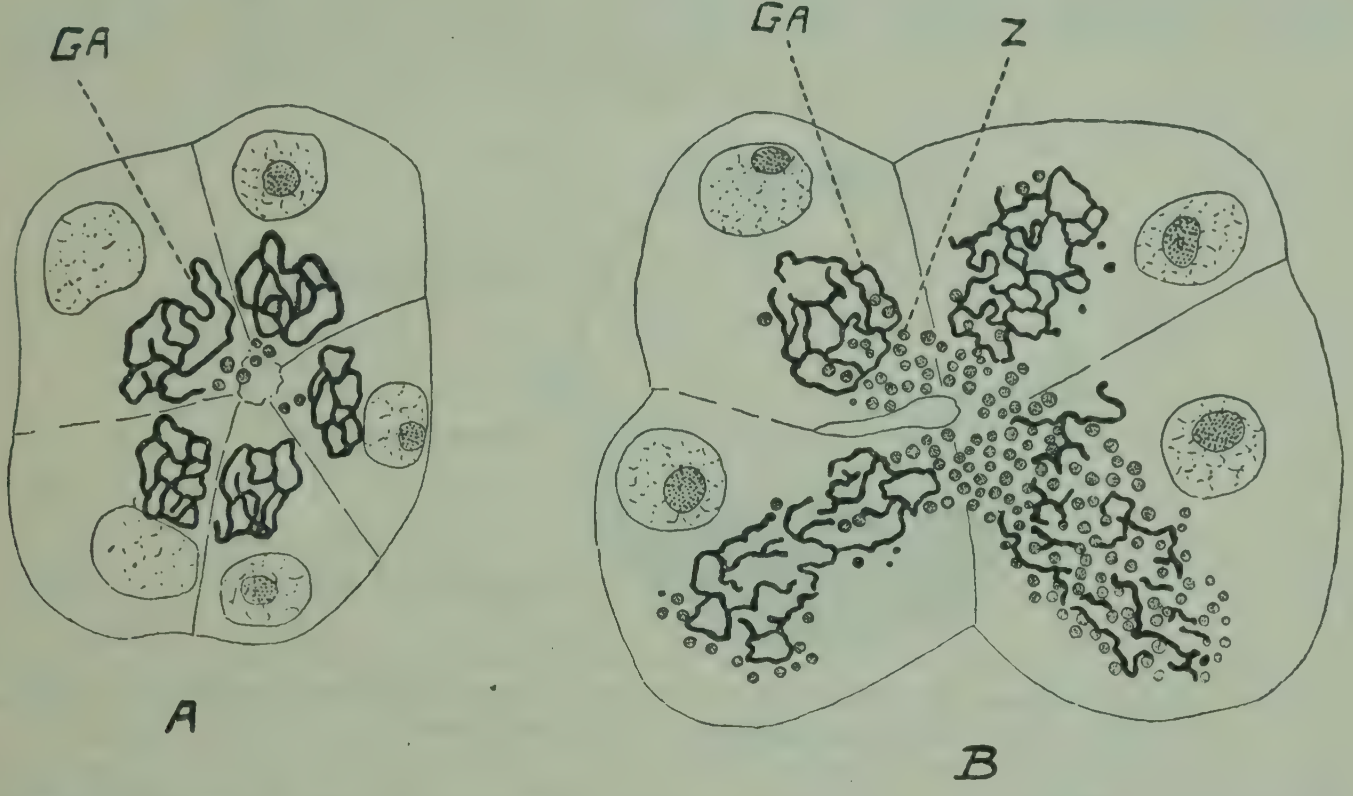


Fig. 6. Diagram of Golgi apparatus (GA) in secreting cells of the pancreas. A. In the contracted condition, after discharge and prior to re-formation of zymogen. B. Zymogen granules (Z) in relation to enlarged form of Golgi apparatus during active formation of zymogen. (Ludford.)

An organ of considerable importance is the *centrosome*, a specialised area of cytoplasm often containing a more deeply staining particle, the *centriole*. In the resting, *i.e.* the non-dividing cell, it commonly lies near the nucleus, and is sometimes double; cells which are incapable of division, *e.g.* nerve cells, have no centrosome, but there may also be no centrosome in certain resting cells which are nevertheless capable on occasion of further division; in those cases, when the time for division approaches, a new centrosome is formed in the cytoplasm. The process of cell-division is apparently heralded and initiated by a preliminary division of the centriole and centrosome. Next there appear birefringent striæ radiating out from each centriole, the structure so formed being called an *aster*. The asters on microdissection behave as relatively rigid elastic bodies, very distinct from the surrounding more liquid cytoplasm, and the two asters grow rapidly in size and appear to repel one another, so that they tend to travel to opposite poles of the cell.

Another cell structure is the so-called *Golgi apparatus*, a reticular structure often located at one pole of the cell, soluble in alcohol, blackened by osmic acid and, after suitable fixation, stainable with silver salts (Fig. 6). In secretory cells it is usually found between the nucleus and the lumen. The functions of the Golgi apparatus are uncertain. Its filaments are of lipoid nature, and it has been suggested that certain products of cell activity, *e.g.* zymogen granules, can be assembled at the surface of the filaments.

---

* Bensley and Hoerr. *Anat. Rec.*, 1934, **60**, 449.

Special inclusions are also met with in certain cells, the best example being the Nissl bodies of nerve cells. Sometimes these structures are artefacts produced in the course of fixation. In addition to these cell organs the cytoplasm may contain visible *granules* of material of various types, which are usually formed by the activity of the cytoplasm or stored there until required. Thus we have granules of *zymogen* found in secreting cells and representing a precursor of the enzymes which the cell liberates during activity, or granules of food material provisionally stored in the cytoplasm until required, *e.g.* fat or glycogen.

THE PHYSICAL STRUCTURE OF PROTOPLASM. It is usually impossible, even under the highest powers of the microscope, to make out any structure whatsoever in living cytoplasm. It is an apparently homogeneous colloidal solution, except for the ectoplasm, which is usually a gel. This gel may be reversibly liquefied, *e.g.* by potassium ions, or by exposure to high hydrostatic pressures (500 atm.): Ca ions make the gel more rigid. This ectoplasmic gelation and liquefaction is probably of importance in bringing about movement of a cell, or within it. Thus, within the pseudopodia in *amœba*, there is an increase of rigidity, which maintains the shape, and this gel is subsequently liquefied. In ciliary motion, cell division, etc., also, the ectoplasm probably plays an important part in causing movement.

In the remaining cytoplasm, or endoplasm, there is also some potentiality for gelation, though more feeble. Various inclusions and granules may be visible within it, but these are not essential components, and are not necessarily living. Differentiation between a more or less regular meshwork and a more fluid portion filling up its interstices, is the rule when cells have been fixed by hardening fluids. All the results obtained in this manner must be regarded with some suspicion, since, as has been shown by Fischer and by Hardy,* it is possible to imitate artificially the various structures, which have been assigned as characteristic of protoplasm, by hardening a homogeneous colloidal solution such as egg-white by orthodox methods. The results of differential staining, especially after fixation, also need further caution in their interpretation, since we know so little of the factors which combine to give any particular result. Thus various theories of protoplasmic structure, which have been put forward as the result of observations on hardened tissues, possess little more than historic interest.

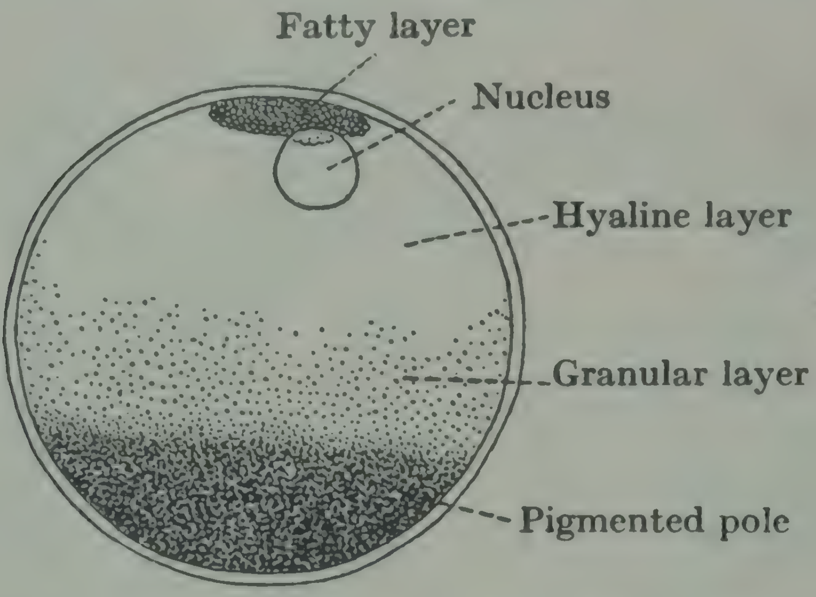


FIG. 7. Stratification in contents of *Arbacia* egg after two minutes' rapid centrifugalisation. (After LYON.)

The perfect potential mobility of the protoplasm of many cells, as instanced by the flow of the substance of an amœba into its pseudopodia, or the ready movements of granules or mitochondria within living cells, indicates a fluid character for the protoplasm. This view is supported by the observation that when small amounts of water are injected into the substance of a cell the water at once mixes freely with the cytoplasm.

Against such a character has been urged the fact that cells have shape, organisation, and power of resistance to deformation—qualities which are generally associated with the possession of solidity. It must be remembered, however, that the ectoplasm is not only often in the form of a gel, but that the whole cell surface must be in a condition of tension, and exercise pressure on the internal portions of the cell, which will tend to bring it into the spherical form. The smaller the mass of protoplasm the greater will be the pressure exerted by its surface layer on its contents and the greater resistance will it present to deformation of the spherical form. A fluid drop, if suspended in a fluid with which it is immiscible, will present greater rigidity the smaller its dimensions. Actual measurement of the tension at the surface of cells, however, shows it to be surprisingly low, viz. 0·1 to 2·0 dynes/cm. This is believed to be due to adsorption of protein on the surface. The endoplasm may therefore be regarded as usually fluid, the form and rigidity which are acquired by most cells being due to physical differentiation in its substance. In some cells, as for instance, muscle fibres, the cytoplasm on microdissection behaves as a gel. The cytoplasm in many cells is certainly of very low viscosity; when certain sea-urchin eggs are centrifuged, the granules settle in layers within the cell (Fig. 7), and in some species at a rate indicating a viscosity only four times that of water. The physical condition of protoplasm is without doubt subject to considerable and often very sudden changes. When, for instance, a living cell is damaged by a microdissection needle, the fluid homogeneous cytoplasm almost at once sets to a jelly-like semi-solid, in which definite granules soon appear. Bayliss showed that when an amœba was stimulated electrically, while observed by darkground illumination, the fine movement of the granules was momentarily checked, also presumably owing to a temporary gelation of the cytoplasm.

---

* HARDY, *J. Physiol.*, 1899, **24**, 158,

When subjected to stress, cytoplasm may show fine fibrillations, which vanish again when the stress is removed. The radiations of the aster are possibly the result of such a state of strain originating in the neighbourhood of the centrosome. They disappear at once if the cell is etherised.

Since the study of the behaviour of the cell shows that it must possess a much more complex structure than that which is revealed by the microscope, and one which permits of the spatial differentiation of the different chemical processes that may occur at one and the same time in the protoplasm, many theories of an ultramicroscopic cell structure have been put forward. Ultramicroscopic examination of living cells certainly shows that the cytoplasm is crowded with granules of various sizes, even though it may appear homogeneous to ordinary microscopic observation. These granules are able to move about within the cell, and indeed, in certain cases, display a lively Brownian movement, *e.g.* in muscle cells, leucocytes, etc. In many animal cells, however, the cytoplasm appears to be too viscous to allow of very active movement.

THE PLASMA MEMBRANE OF CELLS.* Colloid solutions have at their surfaces a layer in which the colloid substance is much more concentrated than it is in the body of the solution. If for instance we take a beaker containing egg-white diluted 100 times, and drop a steel magnetised needle on to the surface, it will float in consequence of the resistance of the surface. If the needle be greasy the same thing will occur on water, and in this case the needle will lie N. and S. On the albumin solution, however, the needle will lie in the position in which it has been dropped. The aggregation of the albumin molecules on the surface of the fluid results in the formation of a solid film which resists any turning of the needle. What applies to the interfacial surface between fluid and air applies also to that between two different fluids. Thus in milk, which is an emulsion of fat droplets in a protein solution, each droplet of fat is surrounded by a layer of concentrated, possibly solid, protein. A similar condensation leading to the formation of a very thin membrane is believed to take place at the surfaces of cells, so that, although the cytoplasm may be liquid, and miscible with water, the contents of the cell are separated from their surroundings by a semi-solid membrane.

Since it is by means of its surface layer that the organism enters into relation with its environment, this layer acquires a prime importance for the life of the cell, and we may therefore consider here at greater length some of the properties of this layer, the *plasma membrane*, as it has been called. This layer is not to be confounded with the cell wall of plant cells, which is generally freely permeable to all kinds of substances. Nor can it be related to any definite histologically demonstrable membrane. The retention of an individuality by the cell must be determined by chemical and physical differences between this layer and the aqueous phases inside and outside it, and one must assume the existence of a definite boundary layer of the protoplasm, even where it is impossible to see any differentiation between this layer and the deeper parts. Investigations into the nature of this frontier between cell contents and environment lead to the conclusion that the plasma membrane is a layer, less than 10 mμ thick, of peculiar molecular structure and remarkable physico-chemical properties. Its electrical resistance is high, and it exhibits unusual permeabilities.

A living cell must take up its food material in the form of solution; and it is the permeability of the superficial layers which will determine the passage of food substances from the surrounding medium into the body of the cell. The immiscibility of the protoplasm with the surrounding fluid shows that the permeability of the membrane must be a limited one. The qualitative aspect of permeability was first studied in vegetable cells. These present within a rigid, but freely permeable, cellulose wall a thin layer of protoplasm (the "primordial utricle"), enclosing a cell sap. The outermost (but invisible) layer of the primordial utricle is the plasma membrane. If the root hairs of tradescantia be immersed in a 10 per cent. solution of glucose or in a 2 to 3 per cent. solution of salt, a process of *plasmolysis* takes place. The cell sap diminishes in amount by the diffusion of water outwards so that the primordial utricle shrinks (Fig. 8 (4)). On placing the cells in distilled water, water passes into the cell sap until the further expansion of the protoplasmic layer is prevented by the tension of the surrounding cell walls (Fig. 8 (1)). This behaviour can be explained only on the assumption that the protoplasm (or its outer layer) is impermeable or but slightly permeable, both to sugar and to salt, but is freely permeable to molecules of water, *i.e.* it behaves as a semi-permeable membrane.

---

* Newton Harvey and Danielli. *Biol. Rev.*, 1938, **13**, 319.

Similar experiments can be made on red blood corpuscles. These also shrink when immersed in salt solutions with a greater molecular concentration than would correspond to the plasma, whereas if placed in weak salt solutions or in distilled water they swell up and burst, discharging their hæmoglobin in solution into the surrounding fluid (*hæmolysis*). By comparison of various solutions it is found that the strength of each solution which is just necessary to cause plasmolysis or hæmolysis, as the case may be, is determined entirely by its osmotic pressure, *i.e.*, solutions having the same osmotic pressures produce the same results. These experiments show the plasma membrane to be permeable to water, but less so to solutes.

The specialised nature of the cell surface is well illustrated by comparing the ionic composition of the cell contents with that of its normal environment. Thus in red blood cells there is a different ratio of K : Na to that in the plasma. This is in part due to the fact that the membrane is impermeable to Na and only slightly permeable to K. Even to the most easily permeable ions, however, the cell membrane is only penetrable with some difficulty, and it has been stated that under the most favourable circumstances the ion permeability of the membrane is only 10^{-8} of that of a layer of water of equal thickness.

The impermeability of the plasma membrane does not apply to all dissolved substances. Overton found that, whereas this layer is practically impermeable to salts, sugars and amino-acids, it permits the easy passage of monatomic alcohols, aldehydes, alkaloids, etc. All these substances are more soluble in ether, oil and similar media than they are in water. The passage of dissolved substances through a membrane may be supposed to depend on the solubility of these substances in the membrane. A layer of lipoid material one molecule in thickness has been shown to hinder the rate of evaporation of water, but not that of ether.* Overton concluded that the superficial layer of cells must itself partake of a "lipoid" character, and that cholesterol and lecithin probably enter largely into its composition. He stated that only such dyes, or other substances, as are soluble in a mixture of melted lecithin and cholesterol have the property of penetrating the living cell, and only those dyes, such as methylene blue, and neutral red, can be used for *intra vitam* staining. This does not apply, however, universally. For example, methylene green, thionin and malachite green are not "lipoid-soluble," but penetrate cells readily, whereas cyanosin and Bengal rose are "lipoid-soluble," but will not enter. The kidney cells can take up many dyes which are not "lipoid-soluble," such as phenol sulphone-phthalein and indigo-carmine, and the same applies to other gland cells.

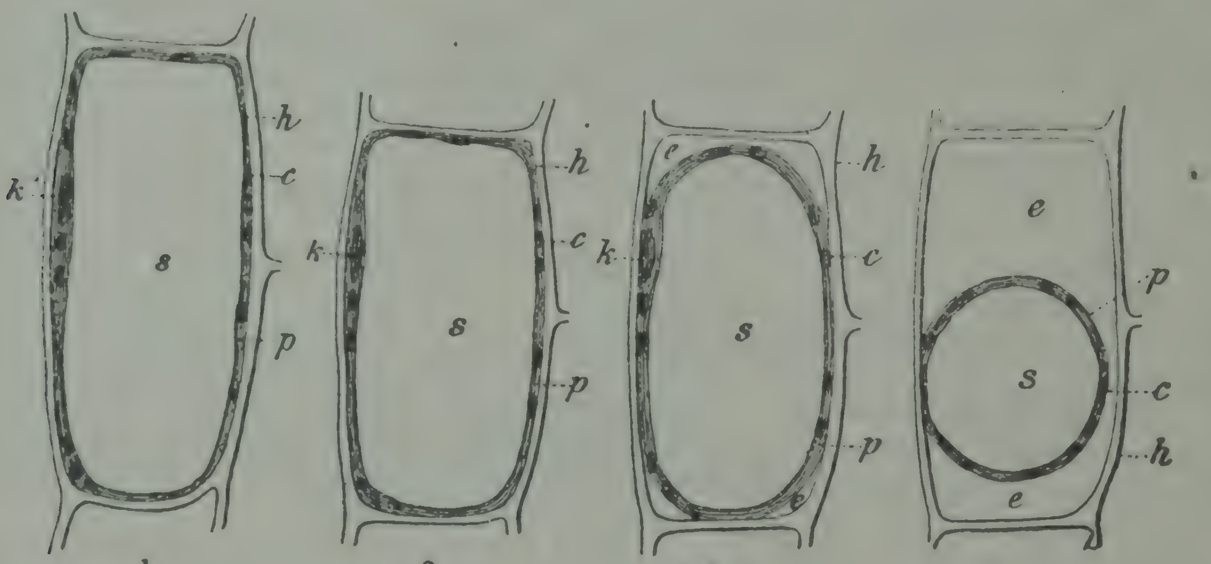


FIG. 8. Vegetable Cells, showing varying degrees of Plasmolysis. (DE VRIES.)

Substances which have the power of dissolving lecithin and cholesterol, such as ether or bile salts, also act as hæmolytic agents, *i.e.* they cause a destruction of the red blood cells by rendering porous the superficial layer which is necessary for their preservation from the solvent effects of the surrounding fluid. Further evidence in favour of the lipoid character of the cell surface is that oil readily wets and adheres to some cells, *e.g.* erythrocytes, *Amœba*; in naked *Arbacia* eggs, an oil globule not only adheres to the surface, but may later pass right through the membrane and collect on its inner surface. For penetration to occur in this mass fashion it has been supposed that the ectoplasmic layer must first be liquefied. An interesting feature of permeability is that weak acids or bases, *e.g.* CO_2, H_2S, HCN, NH_3, penetrate cells readily, whereas strong ones, *e.g.* HCl, NaOH, fail to do so. The semi-permeability of the plasma skin can be altered by changes in the surrounding medium; it seems certain that in some way its integrity is dependent on the presence of certain salts or ions in the outer fluid which may be adsorbed on the external surface of the membrane and so contribute to its qualities. Thus many cells, such as egg cells, muscle cells, entirely alter their permeability if immersed for a time in a fluid free from Ca salts, and under these circumstances may become permeable to substances such as KCl, which they do not ordinarily allow to pass.

---

* RIDEAL. *J. Physical Chem.*, 1925, **29**, 1585. LANGMUIR. *Ibid.*, 1927, **31**, 1719.

It is possible that in order to serve the nutrient needs of the cells, changes may take place in the permeability of the surface layer under certain conditions. Further, in the case of the red blood cells, if we assume complete and permanent impermeability to cations, it is difficult to explain how the relative concentrations of potassium and sodium in cell and plasma could arise in the first place. When cells are damaged, or dead, and often also when they are stimulated, a change in permeability occurs, so that cell contents then diffuse freely through into the surrounding medium.

No doubt protein also enters into the composition of the surface membrane and considerably modifies the properties attributable to lipoids. According to this view * (Fig. 9) the cell surface consists of a solid layer of lipoid perhaps only two or three molecules in thickness, but with a film of protein adsorbed upon its surfaces. The lipoid molecules, at both surfaces, would be arranged perpendicular to the surface like a palisade, with the hydrated polar groups in the oil-water interphase (negative pole nearest the water). The protein is probably either adsorbed to the surface in a partly denatured form, in which case the surface would be wettable with oil, or else each protein molecule would be surrounded by a hydrated envelope ; in that case the surface would be oil-repellant, as is the case with leucocytes.

Some of the permeability properties of the plasma membrane suggest that the surface may be like a mosaic, some patches being lipoid and others not. This really means that there are some properties that are not compatible with the view that the membrane is entirely lipoid. Lipoid-soluble substances would easily penetrate the membrane and would also tend to dissolve and weaken it.

Experiments carried out by micro-dissection methods with living cells demonstrate the elastic and semi-solid nature of the peripheral layers of cells, and show that when the surface of a cell is torn or punctured, or even when part of the cell is cut away, the surface film is often at once reformed. It thus appears to be in many cases a condensation layer formed at the surface of separation in accordance with the laws of surface tension. The presence of calcium appears to be essential for the formation of this membrane.

In such animal cells as are connected together by protoplasmic bridges, *e.g.* the cells of the deeper layers of the epidermis, Chambers found that when one of the cells in the mass was damaged, the coagulative process spreads to neighbouring cells.

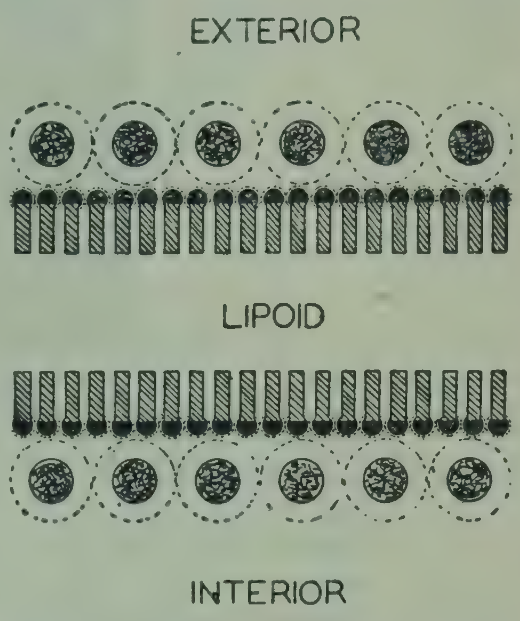


FIG. 9. Schema of molecular conditions at the cell surface. The thin lipoid membrane has its molecules oriented. Outside them is an adsorbed layer of hydrated protein molecules. (DANIELLI and DAVSON.)

THE NUCLEUS. The different staining reactions of nucleus and cytoplasm, and the fact that the nucleus is more opaque to ultra-violet light, suggest a difference in their chemical composition, a suggestion which is confirmed by analysis. Whereas in the cytoplasm the basis consists of ordinary proteins, in the nucleus the chief constituents are nucleo-proteins. Material of somewhat similar nature is occasionally found in the cytoplasm, where it forms *chromidial substance, e.g.* the Nissl bodies of nerve cells or the basal filaments of gland cells such as the pancreas.

The readily stainable material of the nucleus, called *chromatin*, consists mainly of nucleo-protein and "chromosomin."† It can be distinguished from other stainable cell contents by Feulgen's "nucleal" reaction, *i.e.* by treating sections with warm dilute hydrochloric acid for a few minutes, transferring to fuchsin-sulphurous acid (Schiff's reagent) and washing in dilute sulphurous acid. Chromatin is stained violet owing to the presence of pentose in the nucleic acid.

Microscopically, even with dark-ground illumination, the resting nuclear contents are quite homogeneous, except for the presence of nucleoli (Fig. 10). The consistency of the nucleus is found on micro-dissection to be liquid, though perhaps less so than the cytoplasm ; the nucleoli can be freely rolled about within it. When punctured it often rapidly coagulates, and then there may appear, especially in cells about to divide, the characteristic chromatin filaments. These filaments can be drawn out like elastic by the needle, and retract again when released (Fig. 11). Alternatively puncture of the nucleus may lead to disintegration of both nucleus and cytoplasm (Figs. 10, 12).

The nucleus is essential for the continued existence of most cells. By micro-dissection it is possible to divide the larger protozoa into two pieces, one with and one without a nucleus. Hofer, experimenting on the amœba, found that the fragment containing the nucleus quickly regenerated the missing part and pursued a normal existence. On the other hand, the non-nucleated fragments showed no signs of regeneration (Fig. 13). Their movements continued for a short time and then ceased, though the pulsations of the contractile vesicle were but little affected. The power of digestion of food was completely

---

* DANIELLI and DAVSON. *J. Cell. and Comp. Physiol.*, 1935, **5**, 495.

† STEDMAN and STEDMAN. *Nature*, 1944, **153**, 500.

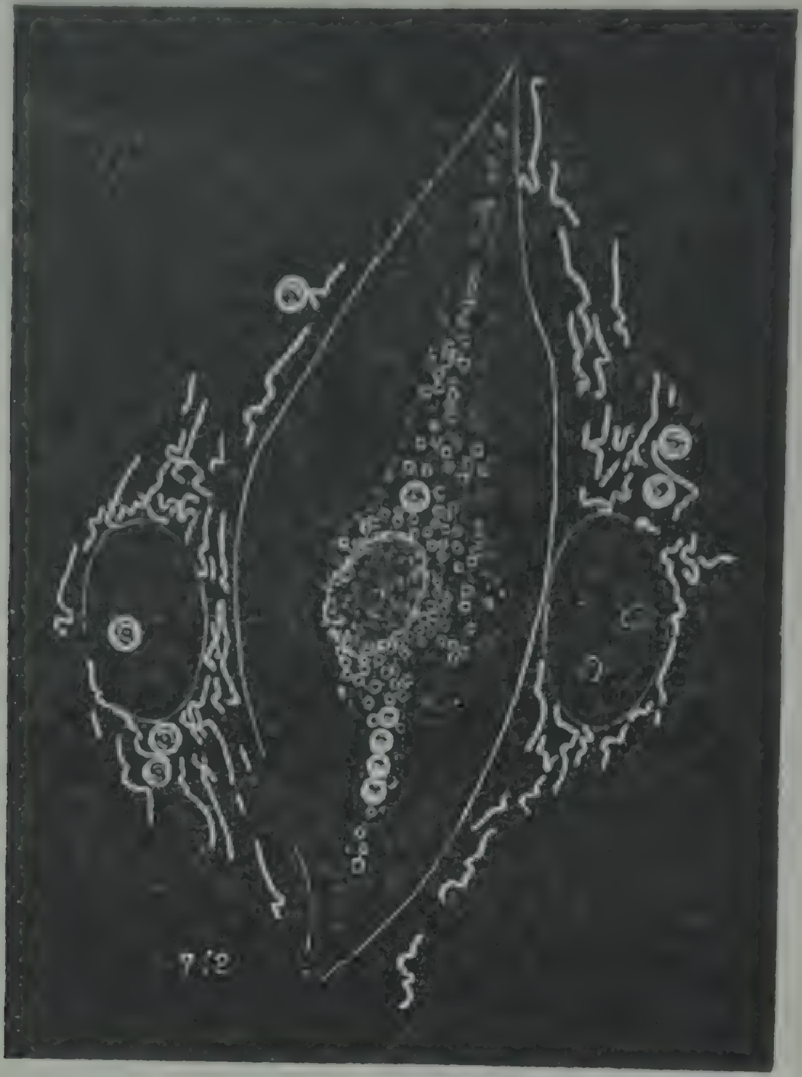


FIG. 10. Three living intestinal epithelial cells in tissue culture, viewed by dark-ground illumination. The nucleus of the middle cell has been punctured by a micro-dissection needle eight minutes previously. The outer boundaries of the two outer cells are invisible, and these cells with their nuclei and mitochondria appear normal. The cytoplasm of the middle cell has shrunk away from its neighbours and is granular: the mitochondria are also altered. (CHAMBERS and FELL. *Proc. Roy. Soc.*, 1931, **109B**, 380.)

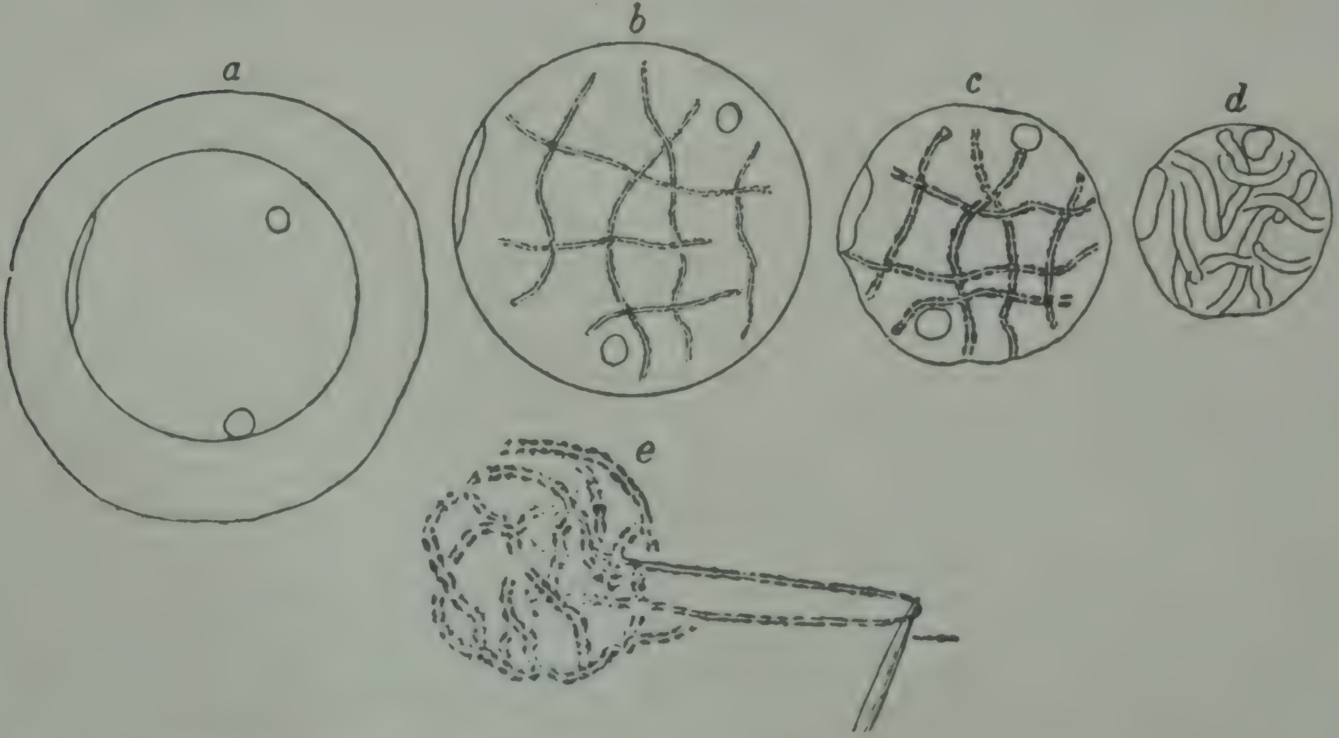


FIG. 11. Living Grasshopper spermatocyte. (*a*) Uninjured; (*b*), (*c*), (*d*) successive stages after injury; (*e*) loop of chromatin filaments stretched with needle. (CHAMBERS.)

lost. In ciliated mammalian cells, the ciliary movement ceases when the nucleus is destroyed or removed. Fibroblasts in tissue culture often have two nuclei. If the nucleus of a uninucleated cell is punctured, the cell dies, but if one nucleus of a binucleated cell is punctured it continues to live.*

The nucleus plays an important part in cell-division. But it is apparently not indispensable for this. When *Arbacia* eggs are centrifuged they elongate, and when the length of such a cell reaches π times its diameter, it breaks into two, one half containing the nucleus. If the non-nucleated half is now subjected to the action of parthenogenetic reagents, it proceeds to divide as a normal cell would, and the subdivision continues until a blastula of about 500 cells is reached, before it finally dies.*

---

* CHAMBERS and FELL. *Proc. Roy. Soc.*, 1931, **109B**, 381.

The nuclear chromatin contains, or consists of, those large nucleoprotein molecules, or micellæ, the *genes*, on which the phenomena of inheritance depend. The transmission of the paternal qualities from one generation to the next is effected by the entrance simply of the nuclear material of the male cell, the spermatozoon, into the ovum.

Conclusions regarding the importance of the nucleus may be drawn from a study of the changes in the nucleus which accompany different phases in the activity of the whole cell. The active state of cytoplasm, which accompanies the activity of secreting cells, is always associated with changes in the position and in the size of the nucleus.

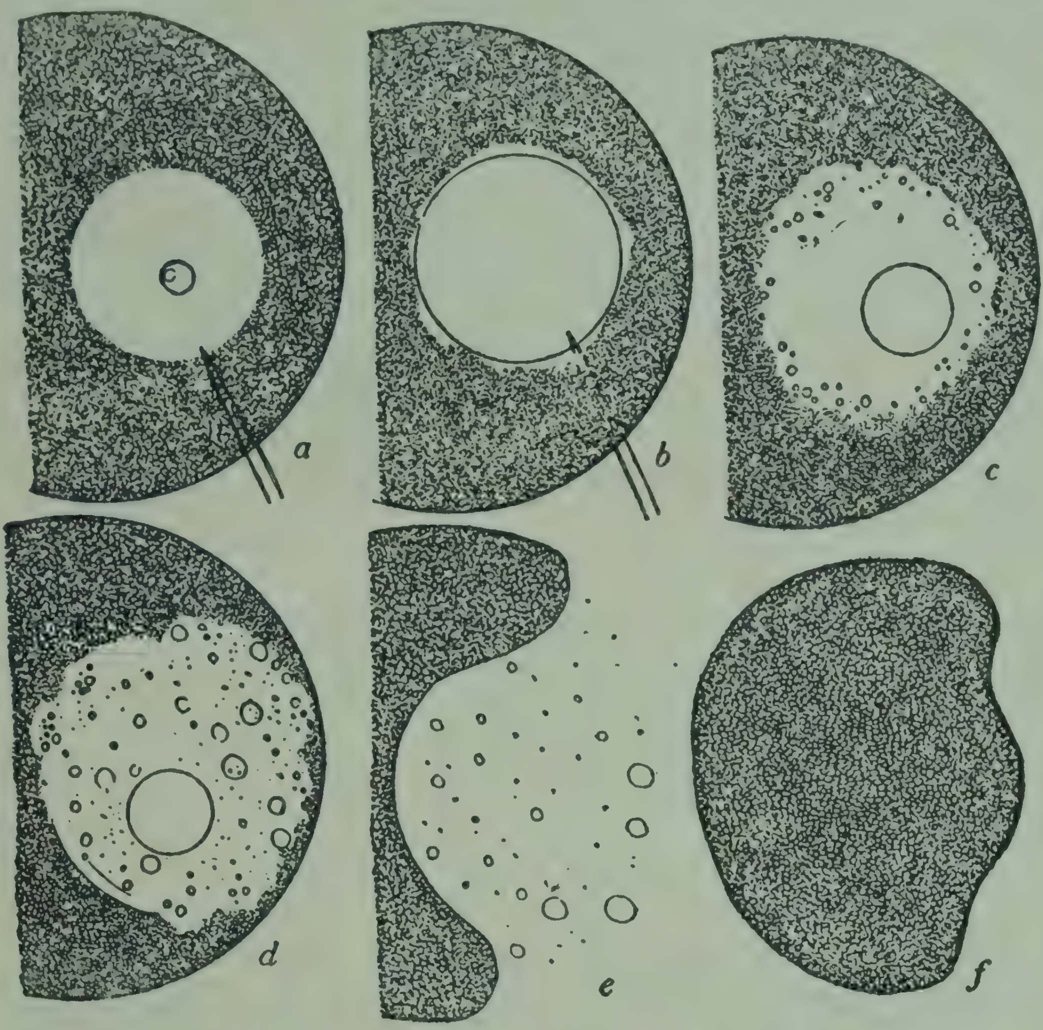


FIG. 12. Disintegration of cytoplasm of starfish egg following puncture of nucleus by micro-dissecting needle. (CHAMBERS.)

HISTOLOGICAL DIFFERENTIATION OF CELLS. Even within the limits of a single cell, differentiation of structure can take place by the setting apart of distinct portions of the cell for isolated functions.

The organs which are thus formed in unicellular animals or plants can be divided into two classes, namely (1) temporary organs, which are formed out of a common structural basis and can therefore be replaced at any time by the cytoplasm if destroyed. Examples of such organs are the cilia of unicellular organisms, and the mouth of animals such as vorticella. In contradistinction to these organs are (2) a number of others which must be regarded as permanent. These cannot be formed by differentiation from the cytoplasm of the cell, but are derived by the division of pre-existing organs of the same character, and are therefore transmitted from one generation to another. As examples of such cell organs may perhaps be mentioned the nucleus with its chromosomes, and the mitochondria and plastids.

The possibility of histological differentiation and of the adaptation of structure to definite functions becomes much more pronounced as we pass to the multicellular organisms. The lowest of the Metazoa, such as the sponges, consists of little more than an aggregation or colony of cells. In the Cœlenterata the differentiation is already much more marked. In all these classes of animals the external medium of every cell forming the organism is the sea-water or other medium in which it lives, and every cell is therefore exposed to all the possible variations which may occur in the composition of the surrounding medium.

A great step in evolution was accomplished with the formation of the Cœlomata, the class to which all the higher animals belong. In these, by the formation of a body cavity containing fluid, an internal medium is provided for all the working cells of the body. The composition of this internal medium is maintained constant by the activity of the cells in contact with it, and the stress of sudden changes in the chemical composition of the surrounding medium is borne entirely by the outer protective layer of epiblast cells. These are rendered more or less impermeable by the secretion on their surfaces of a cuticular layer, and only such of the constituents of the surrounding medium are allowed to enter the organis as can be utilised by it. Out of the cœlom is later on formed a circulatory system whic by the circulation of the cœlomic fluid or of blood through the whole body, can procure still more perfect uniformity in the chemical conditions to which every cell is exposed. On of the main results of the differentiation of function and structure is therefore a gradu setting free of the majority of the cells of the body from the influence of variations in th environment.

---

* E. B. HARVEY. *Biol. Bull.*, 1936, **71**, 101.

The differentiation of the cells which compose the organs of the body, is determined i the first place by the different conditions to which they are exposed in virtue of thei positions in the course of embryonic development. All the higher animals may be considered as built in the form of a tube, the external surface of which is modified for th purpose of defence and for adaptation to changes in the environment. From this laye there are developed not only the protective cuticle, but also the organs of motor response, namely, the special senses and the nervous system. The internal surface of the tube is modified for purposes of alimentation. From it are developed all those structures which serve for the digestion of the foodstuffs, for their absorption into the common circulating fluid, for their elaboration after absorption, and their preparation for utilisation by other cells of the body. Between these two surfaces are situated the supporting tissues of the body as well as the muscles. Here also is the cœlom, represented in the higher animals by the pleural, pericardial and peritoneal cavities. The alimentary canal projects for a considerable part of its course into this cœlom, being attached to the body wall only by one side. From the cœlom is also developed the blood vascular system, which maintains a constant circulation of the blood throughout the body. By this differentiation the body becomes divided into a number of organs, each of which is composed of like cells, modified for a common function and bound together by connective tissue, the latter serving also to carry the blood vessels which convey the common medium for the working cells.

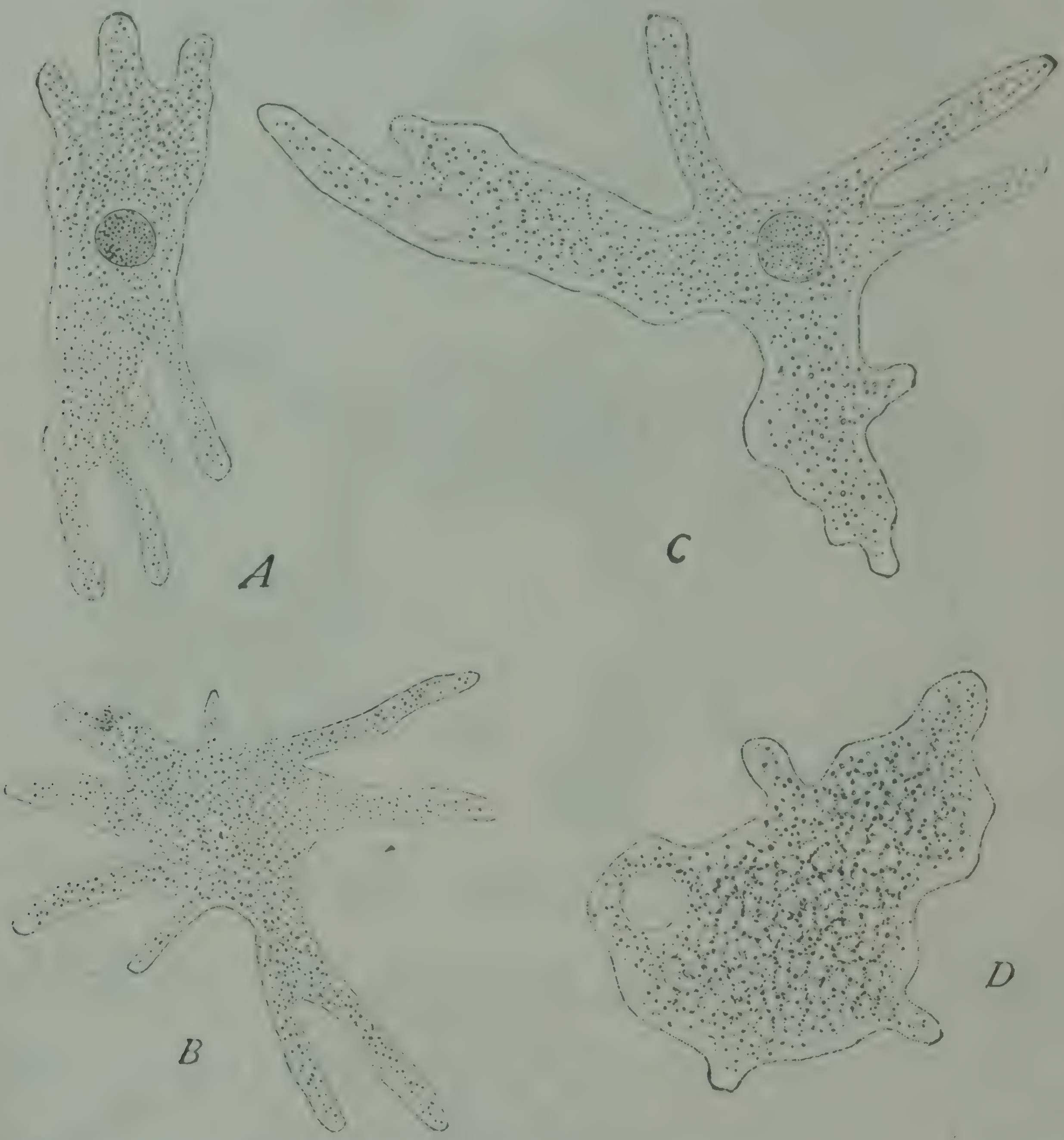


FIG. 13. Nucleated and Non-nucleated Fragments of *Amœba*. (WILSON after HOFER.)
A, *B*. An *Amœba* divided into nucleated and non-nucleated halves, five minutes after the operation.
C. *D*. The two halves after eight days, each containing a contractile vacuole.

The study of the factors which are responsible for the differentiation of the numerous types of tissue cell from the division of the one common cell, the fertilised egg-cell, is a branch of the subject of experimental embryology. It has up to the present not been possible to obtain much beyond a few scattered hints as to what the nature of these factors may be. In the early divisions of some fertilised ova, all the cells appear to be exactly alike. We may believe that to some extent the subsequent course of development of these cells is mechanically determined by their positions and by their relations to one another, but it is certain that this is not the chief factor.

Interesting experiments by Spemann * on embryo newts have shown that at the earliest stages, portions of medullary plate (rudiment of nervous system) and of epidermis can be exchanged, and it is then found that the transplanted portions grow into tissues which are normal for the positions into which they have been transplanted. The position occupied by the tissue is here the decisive factor in its differentiation. But in other experiments it was found that there is a region of the embryo, probably that which later forms the mesoderm, which is not so subjected to its surroundings but which instead determines the formation of definite structures, foreign to their new surroundings. The region is called an "organiser." It seems likely that the directing influence of this region may be due to the presence in its cells of a definite chemical substance, † probably a steroid.

During the course of the development of various organs also it would seem that after a certain stage of embryonic development has been reached, local organising influences are set up. For example, the femur of a five-and-a-half-day chick embryo can be grown *in vitro*, reach more than three times its initial length, and show almost normal form and bone structure ‡ in spite of isolation from its normal surroundings. Similarly, an eye from a three-day chick embryo can be grown *in vitro*, and shows a remarkable development and differentiation into normal eye tissues.§

Experiments in tissue-culture with embryonic tissues also indicate that chemical factors play a part in causing differentiation.

In the foregoing lines we have spoken of an isolated cell of the body as if it were an independent elementary organism. A better term for such an aggregation of cells as presented by the higher animals is not, however, "cell colony," but "cell state," since, just as in the state politic, no cell is independent of the activities of the others, but the autonomy of each is merged into the life of the whole. With increasing differentiation there is increasing division of function among the various members of the state, and each becomes less and less fitted for an independent existence or for the discharge of all its vital functions. Thus the life of the organism is essentially composed of and determined by the reciprocal actions of the single elementary parts. It is evident that, if the process of specialisation has gone far enough, a discussion whether each unit has or has not an independent life is beside the mark, since it cannot possibly exist apart from the activities of the other cells. From this aspect the morphological unit is not the cell but the whole organism, and the division of the common cytoplasm into cells is merely a question of size and convenience. There can be no doubt that an important factor in the division of cells is their growth: the cell divides because it grows. With increased mass of living substance it is necessary to provide for increase of surface both of cytoplasm and of nucleus. Whether all the tissues of the higher animals remain in structural continuity by protoplasmic bridges, etc., must be a matter of indifference, since all that is necessary for the interdependent working of the different cells of the body is a functional continuity, and this in the higher animals is effected by the presence of a common circulating fluid and a reactive nervous system connected with all the cells of the body.

For Reference

Bloom (1927). Cellular Differentiation and Tissue Culture. *Physiol. Rev.*, **17**, 589.
Bourne (Editor) *et al* (1942). Cytology and Cell Physiology. (Oxford. Clarendon Press.)
Carrel (1924). Tissue Culture and Cell Physiology. *Physiol. Rev.*, **4**, 1.
Fischer (1946). The Biology of Tissue Cells. (Cambridge Univ. Press.)
Gersh (1941). Recent Developments in Histochemistry. *Physiol. Rev.*, **21**, 242.
Gray, J. (1931). A Text-Book of Experimental Cytology. (Cambridge.)
Loeb (1930). Transplantation and Individuality. *Physiol. Rev.*, **10**, 547.
Needham (1942). Biochemistry and Morphogenesis. (Cambridge U. P.)
Newton Harvey and Danielli (1938). Properties of the Cell Surface. *Biol. Rev.*, **13**, 3191.
Seifriz (1929). The Structure of Protoplasm. *Biol. Rev.*, **4**, 76.
Spemann, H. (1928). Organizers in Animal Development. *Proc. Roy. Soc.*, **102B**, 177.
Weiss (1935). The So-called Organizer, etc. *Physiol. Rev.*, **15**, 639.
Wilmer (1935). Tissue Culture. (London. Methuen.)

---

* Spemann. *Proc. Roy. Soc.*, 1927, **102B**, 177.
† Waddington *et al.* *Proc. Roy. Soc.*, 1935, **117B**, 289, 310.
‡ Fell and Robison. *Bioch. J.*, 1929, **23**, 767.
§ Strangeways and Fell. *Proc. Roy. Soc.*, 1926, **100B**, 273.

B. BIOCHEMICAL PRINCIPLES *

CHAPTER III

THE ELEMENTARY CONSTITUENTS OF PROTOPLASM

LIVING things contain no chemical elements peculiar to themselves. Every living organism without exception contains the following elements: carbon, hydrogen, oxygen, nitrogen, sulphur, phosphorus, chlorine, potassium, sodium, calcium, magnesium, copper and iron. In addition to these thirteen elements others are found in certain organisms, sometimes to a large extent. Of these elements we may mention especially silicon, iodine, fluorine, bromine, aluminium, nickel, cobalt, zinc, lead, silver, cadmium, lithium, strontium and manganese.† Dealing with the first class, which includes those essential to all forms of life, we find that their relative proportions in living organisms have little or no relation to their proportions in the environment of the organisms. Their presence, however, in the latter is a necessary condition of life. In the case of plants, growth is limited by the amount of the necessary element (*e.g.* potassium, nitrogen, phosphorus) which is present in smallest quantities in the surrounding medium.

Carbon forms the greater part by weight of the solid constituents of living protoplasm and the proximate constituents of living organisms are practically all carbon compounds and hence are combustible. In the inorganic world practically all the carbon occurs as carbon dioxide, which is present to the extent of about 0·04 per cent. in the atmosphere, while vast quantities are immobilised in the form of chalk and limestone. In this condition the carbon is practically removed from the life cycle, the whole of the carbon contained in the tissue of living beings being derived from carbon dioxide present in the atmosphere.

The energy for the conversion of carbon dioxide into the oxidisable forms which make up the tissues of plants and animals, is furnished by the sun's rays. The conversion of the radiant energy into the potential chemical energy of the carbon compounds is performed by the chlorophyll corpuscles in the green parts of plants. In these corpuscles, under the influence of the sun's rays, the carbon dioxide of the atmosphere, together with water, is converted into starch, and the oxygen liberated in the process is set free.

$$\mathrm{n}\,[6CO_2 + 5H_2O] = (C_6H_{10}O_5)_n + 6nO_2.$$

In this process a large amount of energy is absorbed, which can be set free later by the oxidation of the starch to carbon dioxide. In the oxidation of 1 gramme of starch about 4·5 kilo-calories are evolved, and this represents therefore the measure of the solar energy which must be absorbed by the chlorophyll corpuscle in the process of formation of 1 gramme of starch from carbon dioxide. By this means the world of life is provided with a source of energy. At the expense of the energy of the starch further synthetic processes are carried out and fats and proteins manufactured.

All animals take in carbon, hydrogen, nitrogen, oxygen and sulphur in the form of the carbohydrates, fats and proteins which have been built up in the living plants. In the animal organism these foodstuffs serve as sources of energy.

Oxygen is taken up in the free as well as in the combined state. It takes a position apart from the other foodstuffs in that its presence is the essential condition for the utilisation of their potential energy. In the living cells it combines with the oxidisable compounds formed by the agency of the living protoplasm, with the production of carbon dioxide and water and the evolution of energy.

Nitrogen is also derived indirectly from the air. It can be taken up by most plants only in the form of ammonia, nitrites or nitrates. Animals cannot utilise these compounds, and their only source of nitrogen is the protein which has been built up from them by the agency of the plant cell.

Since nitrogen in the free state is useless, the existence of life must depend on the amount of combined nitrogen which is available. Every thunderstorm results in the production of small quantities of ammonium nitrite, which will be washed down with the rain and serve as a source of combined nitrogen to the soil. Many of the higher plants require their nitrogen in the condition of nitrates, and this conversion is effected by micro-organisms. There are a number of bacteria (*bacterium nitrosomonas*) which have the power of converting ammonia into nitrites. Others (*b. nitromonas*) convert nitrites into nitrates. This has been made the basis of a method now largely employed for the treatment of ammoniacal sewage. These different bacteria play an important part in all soils.

---

* Physiology and biochemistry overlap considerably, and it is desirable for the advanced student of physiology to enter more deeply into the biochemical side of the subject than is possible in this volume. For this purpose he may profitably consult:—

BALDWIN (1947). Dynamic Aspects of Biochemistry (Cambridge Univ. Press).

HARROW (1946). Textbook of Biochemistry (Philadelphia and London. W. B. Saunders Co.)

in addition to the various references given in the present volume.

† **Fox and RAMAGE.** *Proc. Roy. Soc.*, 1931, **108B**, 157.

Organisms also exist which can bring into combination the free atmospheric nitrogen itself. Thus certain soils have been found to undergo a gradual enriching in nitrogen although no nitrogenous manure has been applied to them. This fixation of nitrogen by soils is effected by a distinct micro-organism, *clostridium pasteurianum.*

In addition to this spontaneous fixation of nitrogen by humus, it has long been known to farmers that the growth of almost any leguminous crop in a soil poor in nitrogen may result in an actual increase of nitrogen in the soil from which the crop is taken. It was found that this was dependent on the presence on the roots of certain small nodules swarming with bacteria. These bacteria, like those of humus, are able to assimilate the free nitrogen of the atmosphere but, unlike them, can effect this assimilation only in a condition of symbosis, *i.e.* living in the growing tissues of a leguminous plant.

Sulphur is found in all soils in the form of sulphates, generally of calcium. In the plant cell it is built up, together with nitrogen, carbon and hydrogen, to form a part of protein molecules. Practically the whole of the sulphur utilised by animals is in the form of proteins. It is oxidised in the body which it leaves in the form of sulphates.

Boron is essential for the life of some plants.

Iron, although forming but a minute proportion of the material basis of living organisms (the whole body of man contains only six grammes), is nevertheless indispensable for the maintenance of life. It is necessary for instance in two important functions, viz. the formation of chlorophyll in the green plant and the respiratory process in the higher animals. Iron forms an essential constituent of hæmoglobin, the red colouring matter of the blood, whose office it is to carry oxygen from the lungs to the tissues. The pigment chlorocruorin, present in the plasma of polychæte worms, is an iron compound resembling hæmoglobin. It is probable too that the minute traces of iron in protoplasm exercise an important function in the processes of oxidation which are continually going on. Iron obtained by plants from the soil is built up into highly complex organic compounds, and in this form is taken up by animals. It is probable that the main requirements of the animal for iron, which in the day's diet for man is present to the extent of about 15 mg. in all, may be satisfied entirely at the expense of these organic compounds, but there is little doubt that the animal can, if need be, also utilise the iron salts present in its food. The animal proceeds extremely economically with its supply of iron. Any excess of iron above that needed to supply the iron lost to the body is excreted almost entirely with the fæces in the form of sulphide.

Phosphorus is absorbed by the plant as phosphates. It is one of the most important elements and is concerned in many fundamental biochemical processes, usually in the form of organic phosphoric esters of various types. In the cell it is built up with fatty acids and other organic radicals to form complex compounds such as lecithin, and nuclein. The different phosphorus compounds are broken down by processes of dissociation and oxidation, with the production, as a final result, of phosphates which are excreted with the urine or fæces and return to the soil.

Chlorine, potassium, sodium, calcium and *magnesium* are taken up by the plants in the form of salts ; magnesium is one of the elements present in the chlorophyll of plants. They are taken up by animals in the form of salts, and as such are again excreted with the urine.

Little is known about the significance, if any, of the other elements which have been mentioned as frequent constituents of living beings. *Silicon,* which is of very wide distribution, is assimilated as silica, probably in colloidal solution, and is distributed in minute quantities through all plant and animal tissues. It forms a very large percentage of the mineral basis of grasses, but even here it does not seem to be indispensable.

Fluorine is found in the enamel of the teeth and in minute traces in other tissues of the body.*

Bromine, though present in quantity in some seaweeds, appears to play no part in the economy of higher animals.

Iodine is found in large quantities in many seaweeds and sponges. The substance thyroxine, the active principle of the thyroid gland, contains iodine, and is essential for normal life in the higher animals. Iodine therefore is an essential constituent of the higher animals.

Aluminium is found in large quantities in certain lycopods. Whether it is essential to their growth is not known.

Copper in traces is often a valuable constituent of the diet of animals. In the cephalopods and crustacea it appears to take the part of iron in the formation of a blood pigment, hæmocyanin, which plays the same part in the blood of these animals that is played by hæmoglobin in the blood of vertebrates. When oxidised it is of a blue colour, but gives off its oxygen and is reduced to a colourless compound on exposure to a vacuum.

---

* McClure. *Physiol. Rev.*, 1933, **13**, 277.

Manganese is widely distributed in traces, and in the lamellibranch *Pinna squamosa* appears to form an essential constituent of a blood pigment, like hæmocyanin in its functions, and called pinnaglobulin. An ordinary mixed diet for man probably contains about 2 mg. each of copper and manganese per day.

Cobalt is a constituent of the anti-anæmia factor in the liver.

For Reference

IRVING and MANERY (1936). The Significance of the Chlorides in Tissues and Animals. *Biol. Rev.*, **11**, 287.

CHAPTER IV

THE PROXIMATE CONSTITUENTS

In spite of the enormous variety of the proximate organic constituents of living organisms (and hence of all foodstuffs), most of them are members or derivatives of three classes of compounds. These classes are :—

The Lipides.
The Carbohydrates.
The Proteins.

It is assumed that the student has some acquaintance with the elements of organic chemistry and we shall therefore confine our description here to those aspects which have a direct bearing on the physiology of nutrition and metabolism.

1. THE LIPIDES

The ethereal extract of a tissue contains a variety of fat-soluble substances, many of which are characterised by the fact that they contain fatty acids as integral parts of their molecules. They may be classified into the following groups :—

(1) *Fats.* Neutral esters of glycerol and fatty acids.

(2) *Sterols and Waxes.* Of these the sterols are by far the most important. They are alcohols, the chief sterol of the animal body being cholesterol. Waxes are neutral esters of alcohols, such as cholesterol, cetyl alcohol, etc., with fatty acids.

(3) *Phospholipides* (Phosphatides) are nearly insoluble in acetone, a property which is utilised in their separation from other lipides. Lecithin and kephalin are compounds of an alcohol (glycerol) with two fatty acids, phosphoric acid and an organic base (choline or ethanolamine respectively). Sphingomyelin contains phosphoric acid linked to choline and a base, sphingosine, which in turn is linked to a fatty acid.

(4) *Glycolipides* (Galactosides or Cerebrosides). Compounds of a carbohydrate (galactose), with a fatty acid and an organic base (sphingosine).

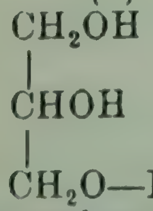
(5) *Hydrocarbons.* Very small amounts of substances belonging to the class of lipochrome pigments occur in the tissues. An important one of these is the hydrocarbon carotene, $C_{40}H_{56}$, which is converted into vitamin A in the liver. Another unsaturated hydrocarbon, squalene, $C_{30}H_{50}$, occurs in very large quantities in some shark liver oils.

1. THE FATS

Quantitatively the fats themselves are easily the most important of the lipide groups and they are widely distributed throughout the animal and vegetable kingdoms. In the higher animals they are the main constituents of the fatty or adipose tissue. They also occur as fine particles in the protoplasm of cells and probably also in combination with the other constituents of protoplasm. In contrast to the mammal, certain fish store up their fat reserve in the liver (*e.g.* cod liver oil), instead of in adipose tissue. Large amounts are also found in certain seeds (cotton seed), and nuts (brazil nut).

COMPOSITION. The fats, the chief representatives of simple lipides, are esters of glycerol and the fatty acids. Glycerol can form esters (glycerides) with one, two or three of its hydroxyl groups. Thus with phosphoric acid the following important compounds are formed :—

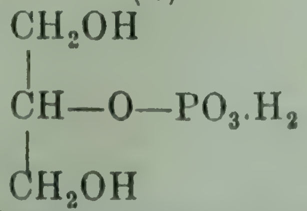
(1)



$$\begin{array}{l} CH_2OH \\ | \\ CHOH \\ | \\ CH_2O{-}PO_3.H_2 \end{array}$$

α glycerophosphoric acid

(2)

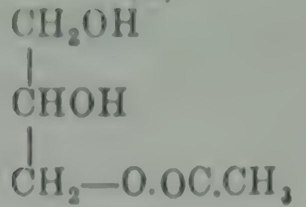


$$\begin{array}{l} CH_2OH \\ | \\ CH{-}O{-}PO_3.H_2 \\ | \\ CH_2OH \end{array}$$

β glycerophosphoric acid

With a fatty acid, such as acetic acid, it forms mono-, di- or triglycerides *e.g.* :—

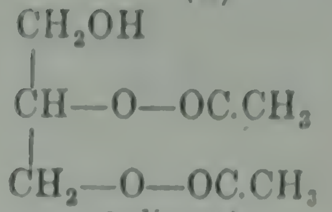
(3)



$$\begin{array}{l} CH_2OH \\ | \\ CHOH \\ | \\ CH_2—O.OC.CH_3 \end{array}$$

α monacetin
(monoglyceride)

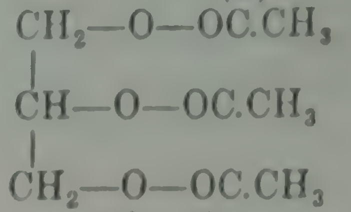
(4)



$$\begin{array}{l} CH_2OH \\ | \\ CH—O—OC.CH_3 \\ | \\ CH_2—O—OC.CH_3 \end{array}$$

α, β diacetin
(diglyceride)

(5)



$$\begin{array}{l} CH_2—O—OC.CH_3 \\ | \\ CH—O—OC.CH_3 \\ | \\ CH_2—O—OC.CH_3 \end{array}$$

triacetin
(triglyceride)

The fats are mainly triglycerides, and natural fat is usually found to consist of various mixed triglycerides ; these are not generally simple esters of the type shown in formula (5), but have different fatty acid radicals combined with the glycerol, and in the animal these fatty acids usually have sixteen or more carbon atoms. The fatty acids which enter into the composition of the triglycerides may be either saturated or unsaturated, belonging to homologous series having the following formulæ :—

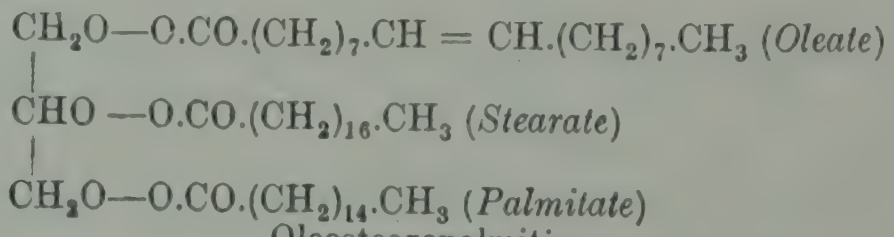
Saturated fatty acids : $C_nH_{2n+1}COOH$ *e.g.*

Butyric acid,	C_3H_7COOH
Palmitic acid,	$C_{15}H_{31}COOH$
Stearic acid,	$C_{17}H_{35}COOH$

Unsaturated fatty acids, e.g. :—

Series $C_nH_{2n-1}COOH$	Oleic acid,	$C_{17}H_{33}COOH$	(1 double bond)
,, $C_nH_{2n-3}COOH$	Linoleic acid,	$C_{17}H_{31}COOH$	(2 double bonds)
,, $C_nH_{2n-5}COOH$	Linolenic acid,	$C_{17}H_{29}COOH$	(3 double bonds)
,, $C_nH_{2n-7}COOH$	Arachidonic acid,	$C_{19}H_{31}COOH$	(4 double bonds)

Adipose tissue, so-called, consists almost exclusively of the fats derived from palmitic, stearic and oleic acids, *e.g.* dioleo stearin, oleo-palmito stearin, triolein, etc., *e.g.* :—



$$\begin{array}{l} CH_2O—O.CO.(CH_2)_7.CH = CH.(CH_2)_7.CH_3 \ (\textit{Oleate}) \\ | \\ CHO —O.CO.(CH_2)_{16}.CH_3 \ (\textit{Stearate}) \\ | \\ CH_2O—O.CO.(CH_2)_{14}.CH_3 \ (\textit{Palmitate}) \end{array}$$

Oleostearopalmitin

The great differences in the appearance of the fat of different animals are due to the varying proportions in which these different mixed glycerides occur. While triolein is liquid at 0° C., tristearin and tripalmitin are solid at the temperature of the body. According, therefore, to the relative amounts of these three fatty acids present in the mixed glycerides, we may have a fat which, like mutton suet, is solid at the body temperature, or a fat containing much olein which is still fluid when the body is opened after death, even when it has already cooled. Milk fat is the one animal fat which contains large proportions of lower fatty acids (butyric acid 5 per cent.).

PROPERTIES. The fats are colourless substances devoid of smell. They float, but are insoluble, in water. They are soluble in warm absolute alcohol, but separate out into crystalline form on cooling, and are easily soluble in ether, petroleum ether, acetone, benzene and chloroform. If they are heated with steam or, more readily, with dilute alkalies, or submitted to the action of certain enzymes, they undergo hydrolysis, and are finally split into three molecules of fatty acid and one molecule of glycerol, *e.g.* :—

$$\underset{\text{(neutral fat—tripalmitin)}}{C_3H_5(OOC.C_{15}H_{31})_3} + 3H_2O = \underset{\text{(palmitic acid)}}{3C_{15}H_{31}.COOH} + \underset{\text{(glycerol)}}{C_3H_5(OH)_3}$$

The action takes place in stages, at each of which one molecule of fatty acid is removed, with the formation, successively, of diglyceride, monoglyceride and glycerol. This process may occur spontaneously when fat is left exposed to the air. Fat is then said to be rancid and has an increased acid value. Most natural fats generally contain a small amount of fatty acid.

Saponification. On boiling a neutral fat for a long time with a solution of potassium or sodium hydroxide, the fat, like other esters, undergoes *saponification*, giving the alkaline salt of a fatty acid, and glycerol. The former compound is spoken of as a soap. In water soap forms a colloidal solution, which, if concentrated, sets to a gel on cooling. From a dilute watery solution the soap can be thrown down in the solid form by the addition of neutral salts.

Emulsification. An emulsion is a system of two immiscible liquid phases, one of which is "dispersed" as globules in the other, as, *e.g.* oil in an aqueous solution, which is spoken of as an oil-in-water emulsion. Milk is such an emulsion, and butter is one of "water" in oil. Oil-in-water emulsions are important, *e.g.* in digestion, since the provision of a large surface area by emulsification of the fat accelerates the action on it of fat-splitting enzymes.

Permanent concentrated emulsions cannot be formed merely by shaking the two components together. In order to form them, at least one other substance, called an *emulsifying agent*, must be present. The emulsifying agent is believed to exert its effect by being adsorbed as a film on the surface of the dispersed globules. The dispersing and stabilising effect of this film is due to (*a*) lowering of interfacial tension between the two phases, thus opposing coalescence of globules; (*b*) establishment of an electric charge at the surface, which also opposes coalescence: and (*c*) a purely mechanical action in coating the globules.

An enormous number of substances, mostly colloids, can act as emulsifying agents under appropriate conditions, the best one in each instance depending on the properties of the substances in the two phases. Thus, soap, fatty acid, bile salts, lecithin and protein can each, under appropriate conditions, act as emulsifying agents for fat/water systems: for example, if we shake together olive oil, soap solution and bile salts, a good emulsion results.

Analysis of Fats. Owing to the difficulty which attends the complete separation of each acid from a mixture of allied fatty acids, such as is obtained by the hydrolysis of a fat, the complete analysis of a sample of a fat is seldom attempted, and in routine work a series of empirical determinations is made which together suffice to give a good deal of information as to the chemical nature of the fatty acids concerned. The following determinations are usual:—

The *melting point.*

The *acid value* of a fat is the amount of potassium hydroxide in milligrammes required to neutralise the free fatty acids in 1 gramme of fat.

The *saponification value* is the amount of potassium hydroxide in milligrammes required to saponify 1 gramme of fat.

The *iodine value* is the amount of iodine in grammes with which 100 grammes of fat will combine.

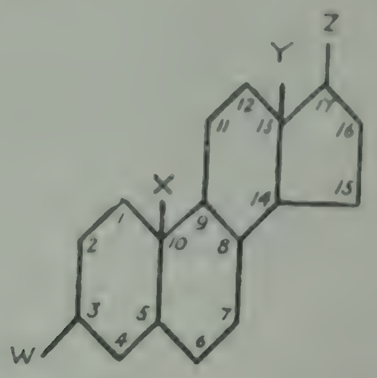
The *Reichert-Wollny value* or *volatile fatty acid value* is the amount of N/10 potassium hydroxide in c.c. required to neutralise the volatile fatty acids in 5 grammes of hydrolysed fat.

The acid value gives an indication of the acidity, or rancidity, of the fat. The saponification value gives some idea of the mean molecular weight of the total fatty acids. The Reichert-Wollny value indicates the proportion of these which are volatile in steam (*i.e.* are of low molecular weight); and the iodine value represents the mean degree of unsaturation of the acids of the fat.

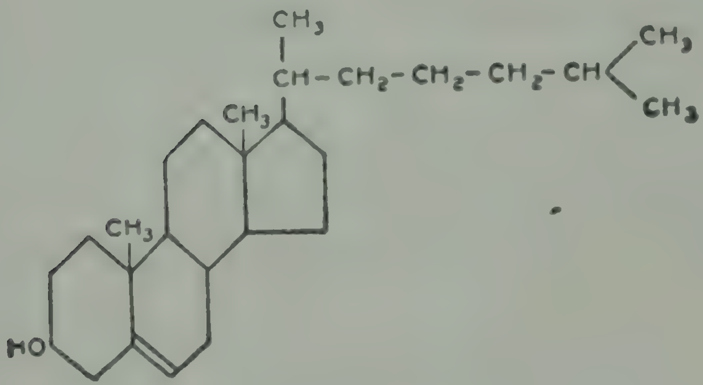
2. THE STEROLS

A certain number of substances occur in the body derived from a formation of esters of the fatty acids with sterols such as cholesterol. The fatty secretion of the sebaceous glands in man and the higher animals, which furnishes the natural oil of hair, wool (lanolin) and feathers, consists of cholesteryl esters with small traces of glycerides. The sterols also occur in the cells in the free state.

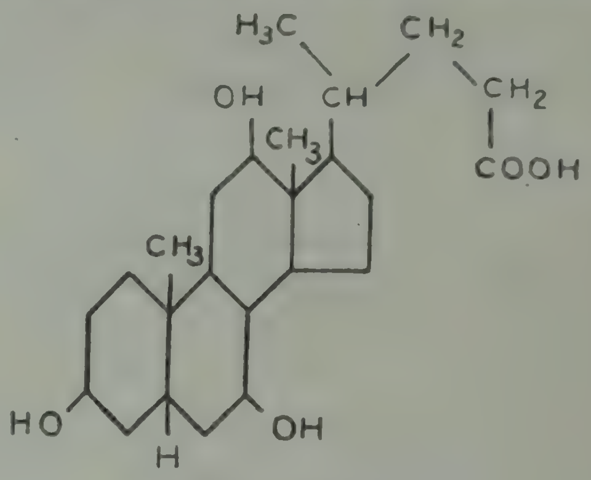
The sterols are mostly complex monatomic alcohols which are widely distributed throughout all the cells and fluids of the body. They contain the characteristic "*cholane*" ring :—



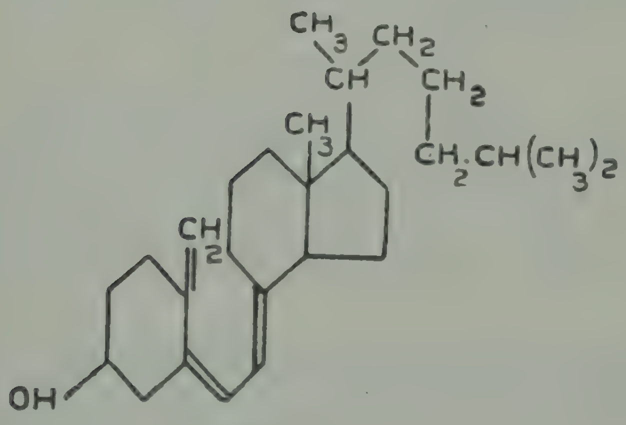
The most widespread member is *cholesterol*, an unsaturated secondary alcohol, $C_{27}H_{45}OH$, which is possibly the precursor in the body of the numerous derivatives. Its corresponding saturated compound, *coprosterol*, is found in fæces,



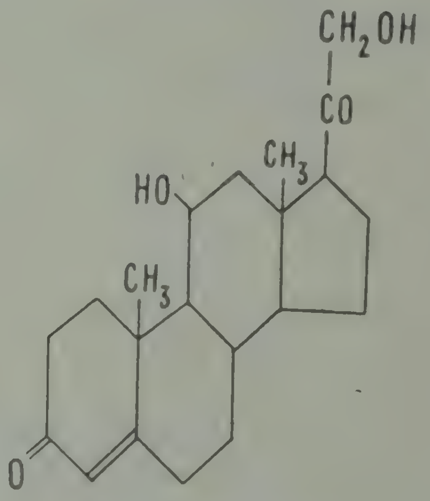
Cholesterol.



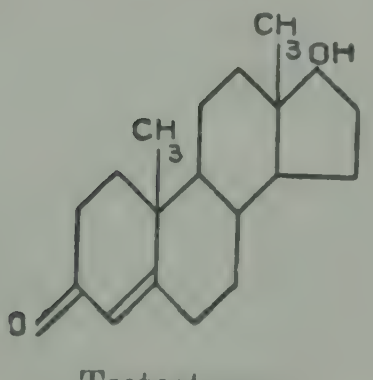
Cholic acid.



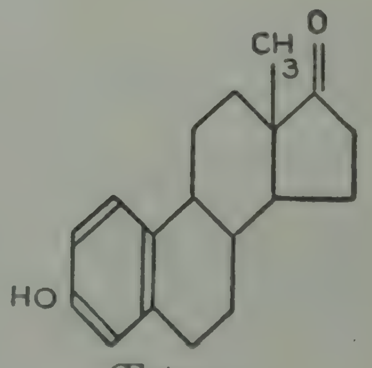
Vitamin D_3.



Corticosterone.



Testosterone.



Œstrone.

while the closely related *ergosterol*, $C_{28}H_{43}OH$, yields one of the vitamins D on irradiation. All the sterols are very stable, are optically active, and resist the action of enzymes and most bacteria.

The sterols are now recognised as a very important class of biological compounds. How closely related chemically are some substances of widely different physiological properties may be seen by the formulæ given above for cholesterol, a bile acid, a vitamin, a suprarenal hormone (corticosterone), a male sex hormone and a female sex hormone (œstrone).

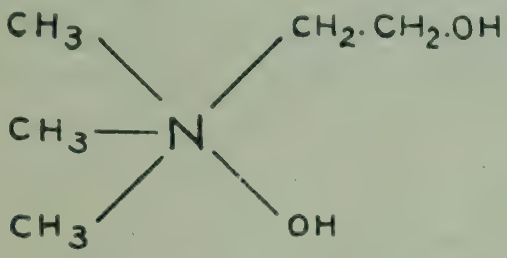
Cholesterol is an important constituent of all cells and it, or its esters, forms part of the framework of the living cells, in the interstices or at the surfaces of which the numerous biochemical reactions associated with their life processes proceed.

For Reference

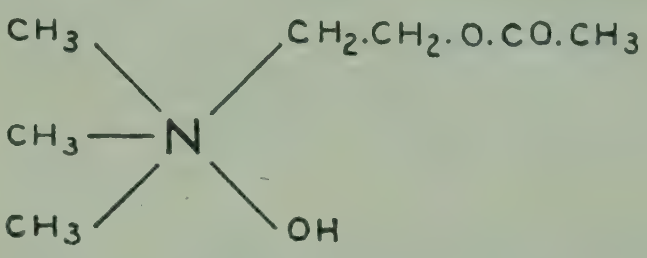
Bills (1935). Physiology of the Sterols. *Physiol. Rev.*, 15, 1.
Friedmann (1937). Sterols and Related Compounds. (Cambridge: Heffer.)
Sobotka (1938). The Chemistry of the Sterols. (London: Baillière, Tindall and Cox.)

3. PHOSPHOLIPIDES

The chief phospholipide is *lecithin*, which contains glycerol, two molecules of fatty acid, phosphoric acid, and the base choline, a substance which has recently acquired considerable importance.



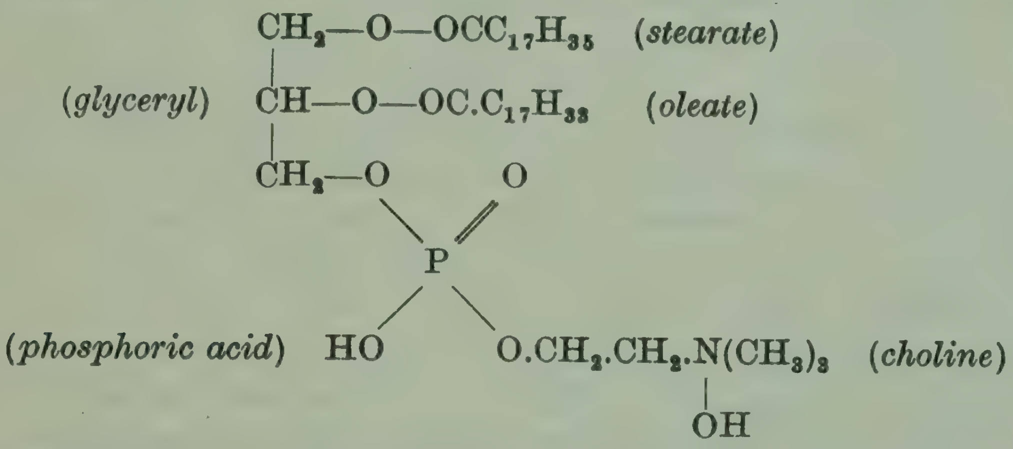
Choline.



Acetyl choline.

Choline has an important biochemical function as a transmethylating agent, which furnishes methyl groups for transfer to other molecules. One of its derivatives, *acetyl choline*, produces potent physiological effects.

Various lecithins, differing only in the nature of the fatty acids contained in them, have been prepared from different tissues. In general, their fatty acids are much more highly unsaturated than those of the food or the reserve fats of the animal—a fact which might suggest that the desaturation of fatty acids is an important preliminary to their metabolism. One of the various lecithins is represented by the following formula in which the two fatty acids are stearic acid and oleic acid :—



Other more unsaturated acids, linoleic, linolenic or arachidonic acids, are however usually present.

Lecithin is a substance very unstable to oxygen, owing to the highly unsaturated acids present in it, and for this reason extreme precautions must be taken in its preparation from tissues. Much of the work on the physiological effect of the phospholipides is probably vitiated by the presence of oxidation products. The universal distribution of lecithin signifies an important function in cellular activity. It undoubtedly plays an important part in the transport and metabolism of fats, related on the one hand to the fact that it is capable of colloidal solution in water, facilitating its free transport and, on the other, to the possibility that desaturation is involved in its production from neutral fats. Some work, however, suggests that the highly unsaturated acids of lecithin may be derived direct from those present in the ingested food. The importance of the constituent base choline in the control of the amount of fat in the liver is discussed later.

It is not improbable that the phosphoric acid of the phospholipides is concerned in the synthesis of the phosphorus-containing nucleo-proteins of the cell, and perhaps also in the normal processes of carbohydrate metabolism. Apart altogether from the chemical significance of its constituents, it is certain that lecithin and the allied substances, by virtue of their special physico-chemical condition in solution, play an important part in determining the peculiar permeabilities, osmotic relations and surface conditions of the cell.

Kephalin has a structure identical with that of lecithin, save that the choline of the latter is replaced by amino ethyl alcohol (ethanolamine), $CH_2NH_2CH_2OH$, linked to the phosphoric acid through its alcoholic hydroxyl group.

4. THE GLYCOLIPIDES

The glycolipides are found chiefly in nervous tissue, and are generally prepared from the acetone extract of brain. They do not contain phosphoric acid. The organic base present is a complex substance called sphingosine, present with a higher fatty acid and galactose. They are relatively more stable to oxygen than is lecithin.

The characterisation of individual conjugated lipides is a matter of much difficulty owing to their instability, to the similarity of their general properties and to their eccentric solubilities.

For Reference

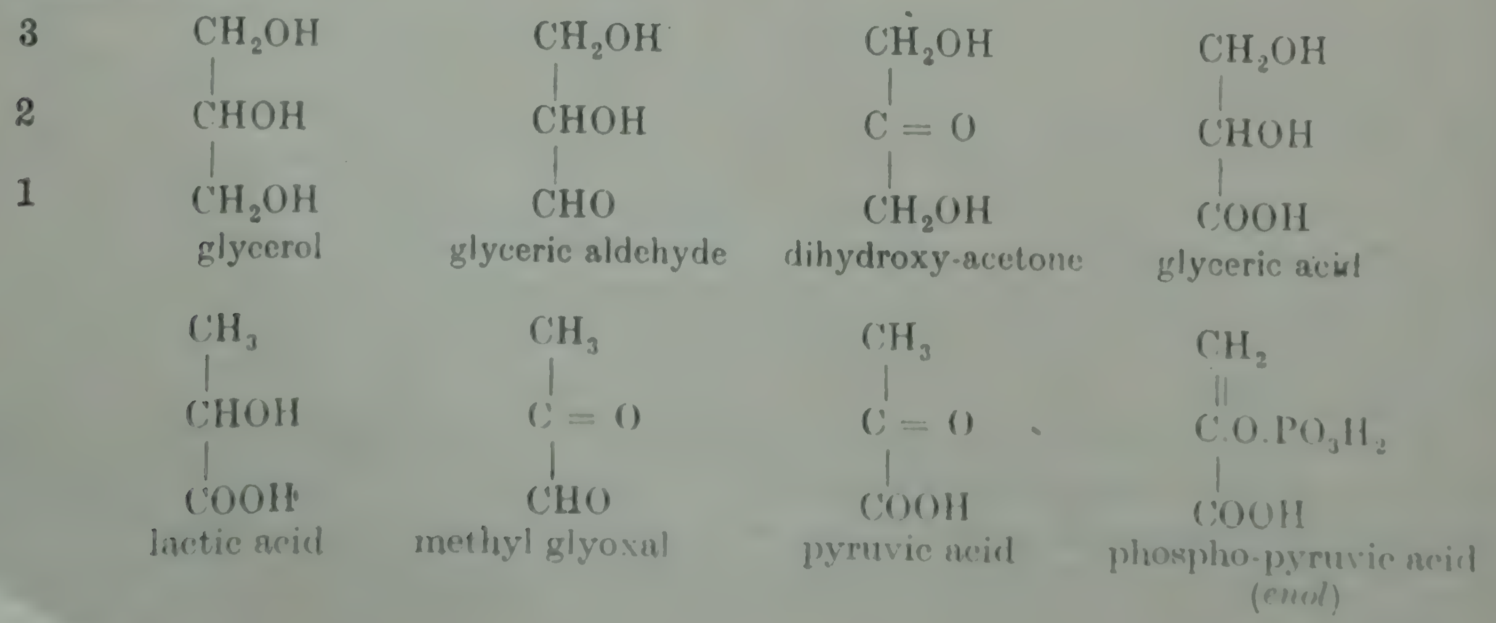
ALLES (1934). The Physiological Significance of Choline Derivatives. *Physiol. Rev.*, **14**, 276
BLOOR (1925). Biochemistry of the Fats. *Chem. Rev.*, **2**, 243.
HILDITCH (1940). The Chemical Constitution of Natural Fats. (London. Chapman and Hall.)
LEATHES and RAPER (1925). The Fats. (Longmans, Green & Co.)
LEVENE (1921). Structure and Significance of the Phosphatides. *Physiol. Rev.*, **1**, 327.
SINCLAIR (1934). The Physiology of the Phospholipoids. *Physiol. Rev.*, **14**, 351.

2. THE CARBOHYDRATES

The Carbohydrates are a group of bodies of wide distribution and great importance in both the vegetable and animal kingdoms. In plants the first product of assimilation of carbon is a carbohydrate, and in animals these substances form one of the most important sources of energy. They consist of the elements carbon, hydrogen and oxygen, the two last-named being almost invariably in the proportions necessary to form water. Their general formula might be expressed $C_nH_{2n}O_n$. Many of the carbohydrates which are of importance in the animal economy contain six carbon atoms or a multiple of this number. Analogous substances, however, can be prepared containing less or more than this number of carbon atoms. A series of compounds exist which contain in their molecule 2, 3, 4, 5, 6, 7, 8, 9 carbon atoms, and are termed dioses, trioses, tetroses, pentoses, hexoses, heptoses, and so on; the termination "ose" with the Greek numeral prefixed, indicating the number of carbon atoms, gives them a distinct designation. They are all either ketones or aldehydes of polyatomic alcohols, and known as *ketoses* and *aldoses* respectively.

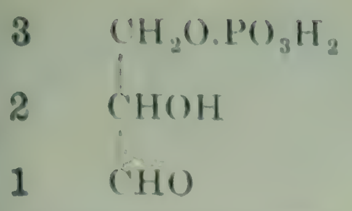
THE TRIOSES ($C_3H_6O_3$)

Two important trioses are glyceric aldehyde (an aldose) and dihydroxy-acetone (a ketose). Both can be obtained by oxidation of glycerol. Their relation to glycerol and to other physiologically important 3-carbon substances are shown below:—



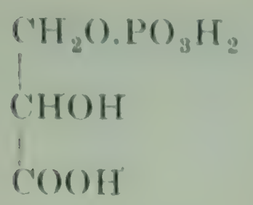
$$\begin{array}{ccccc} 3 & CH_2OH & CH_2OH & CH_2OH & CH_2OH \\ & | & | & | & | \\ 2 & CHOH & CHOH & C=O & CHOH \\ & | & | & | & | \\ 1 & CH_2OH & CHO & CH_2OH & COOH \\ & \text{glycerol} & \text{glyceric aldehyde} & \text{dihydroxy-acetone} & \text{glyceric acid} \end{array}$$

$$\begin{array}{cccc} CH_3 & CH_3 & CH_3 & CH_2 \\ | & | & | & \| \\ CHOH & C=O & C=O & C.O.PO_3H_2 \\ | & | & | & | \\ COOH & CHO & COOH & COOH \\ \text{lactic acid} & \text{methyl glyoxal} & \text{pyruvic acid} & \text{phospho-pyruvic acid} \\ & & & (\textit{enol}) \end{array}$$



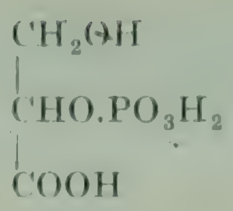
$CH_2O.PO_3H_2-CHOH-CHO$

3-phosphoglyceraldehyde



$CH_2O.PO_3H_2-CHOH-COOH$

3-phosphoglyceric acid

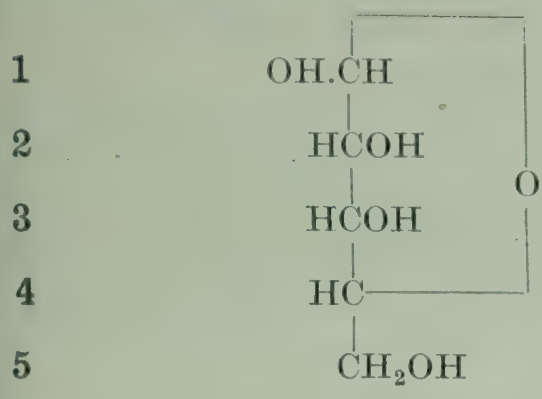


$CH_2OH-CHO.PO_3H_2-COOH$

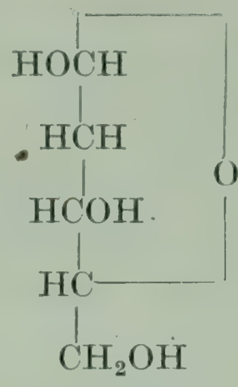
2-phosphoglyceric acid

THE PENTOSES

These bodies occur largely in plants in the form of complex polysaccharides, the Pentosans, which give pentoses on hydrolysis with acids. Three forms of pentose have been found in the animal body, namely, i-arabinose, which has been isolated from the urine in cases of pentosuria, and d-ribofuranose and desoxyribofuranose,



β-d-ribofuranose

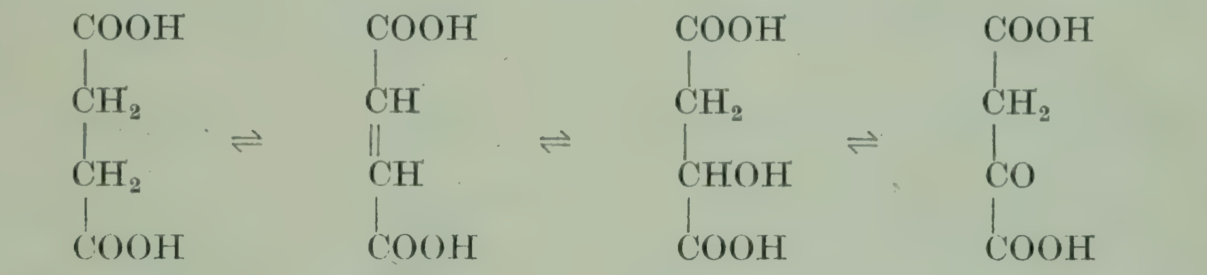


β-d-2 desoxyribofuranose

which occur built up into the nucleic acids, and in various important nucleotides.

SOME IMPORTANT 4- AND 5-CARBON DI- AND TRI-CARBOXYLIC ACIDS

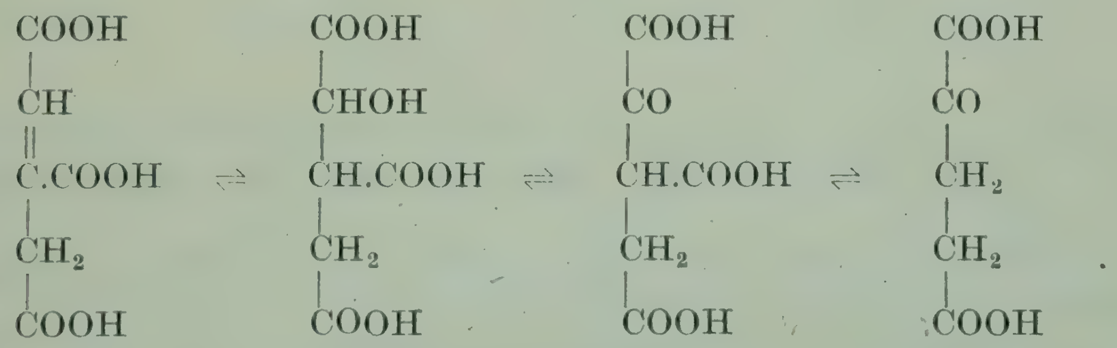
Some compounds which fall under these headings are included here because they probably play an important part in the metabolism of carbohydrates. Four of these 4-carbon dicarboxylic acids are shown below ; they are convertible as shown :—



$COOH-CH_2-CH_2-COOH \rightleftharpoons COOH-CH{=}CH-COOH \rightleftharpoons COOH-CH_2-CHOH-COOH \rightleftharpoons COOH-CH_2-CO-COOH$

Succinic acid. *Fumaric acid.* *Malic acid.* *Oxaloacetic acid.*

and four 5-carbon tricarboxylic and dicarboxylic acids, with convertibilities, are :—



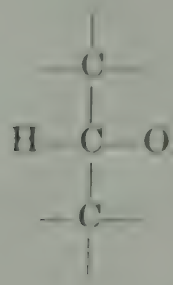
$COOH-CH{=}C.COOH-CH_2-COOH \rightleftharpoons COOH-CHOH-CH.COOH-CH_2-COOH \rightleftharpoons COOH-CO-CH.COOH-CH_2-COOH \rightleftharpoons COOH-CO-CH_2-CH_2-COOH$

Aconitic acid. *Iso-citric acid.* *Oxalosuccinic acid.* *a-ketoglutaric acid.*

THE HEXOSES OR MONOSACCHARIDES ($C_6H_{12}O_6$)

It is convenient to call those compounds containing six carbon atoms the simple sugars, or hexoses.

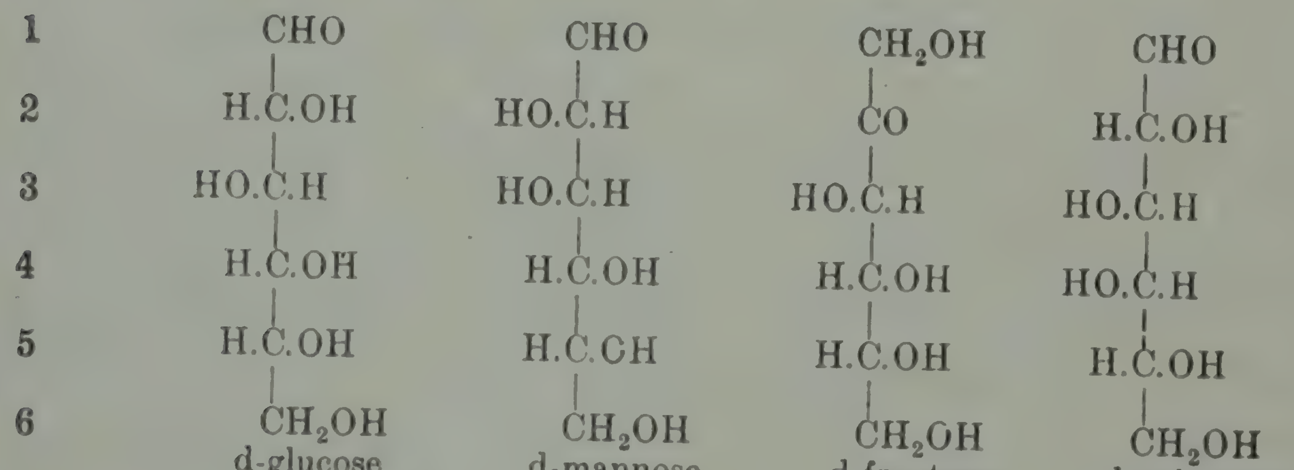
STEREOISOMERISM IN THE SUGARS. It will be noticed that of the six carbon atoms contained in the hexose molecule, *e.g.* the aldose $CHO.(CHOH)_4.CH_2OH$, four are *asymmetric*, *i.e.* their four combining affinities are combined with groups of different kinds, viz. two carbon chains, one H atom, and one OH group



$$\begin{array}{ccc} & | & \\ - & C & - \\ & | & \\ H & -C- & OH \\ & | & \\ - & C & - \\ & | & \end{array}$$

They must therefore present many stereoisomeric optically active forms. Thus an aldohexose with four asymmetric carbon atoms $(CHOH)_4$ must present 2^4 isomers, *i.e.* sixteen isomeric compounds, so that there must be sixteen optically active sugars all possessing the formula $CH_2OH(CHOH)_4CHO$. Only a small number are of any physiological importance. These include the aldoses glucose, mannose and galactose, and the ketose fructose.

Since these sugars can be divided into the dextro-rotatory and lævo-rotatory optically active varieties, an obvious mode of designation would be to represent them as d- or l- varieties respectively. But this mode of nomenclature has been altered in favour of representing by the letter prefixed, not the optical qualities of the substance in question, but its structural relation to dextro-rotatory glucose. Thus d-fructose means that fructose is the ketose corresponding structurally to the dextro-rotatory glucose, d-fructose itself being lævo-rotatory, though its active asymmetric C atoms are identically arranged with those in glucose. With this definition one may say that it is only the d-hexoses of a particular form which are assimilable, and therefore of physiological importance. If the optical activity is indicated, this is done by enclosure in brackets. Thus d (−) fructose means the fructose related to d-glucose, and of lævo-rotatory optical activity. The small differences in the configuration of the four d-sugars can be readily seen if their graphic formulæ, which for simplicity are shown as straight chains, be compared :



$$\begin{array}{ccccc} 1 & CHO & CHO & CH_2OH & CHO \\ 2 & H.C.OH & HO.C.H & CO & H.C.OH \\ 3 & HO.C.H & HO.C.H & HO.C.H & HO.C.H \\ 4 & H.C.OH & H.C.OH & H.C.OH & HO.C.H \\ 5 & H.C.OH & H.C.CH & H.C.OH & H.C.OH \\ 6 & CH_2OH & CH_2OH & CH_2OH & CH_2OH \end{array}$$

d-glucose d-mannose d-fructose d-galactose

Only four hexoses out of the large number which have been synthesised are assimilable by the animal body. These are mannose, glucose, galactose and fructose. They may be synthesised in several ways. The most interesting, because it probably represents the mechanism of synthesis of hexoses in plants, is the formation from formaldehyde.

d-(+)-GLUCOSE is the chief constituent of the sugar of fruits. It occurs in the body as the end product of the digestion of starch. It forms white crystals which melt at 86° C. and lose the one molecule of water of crystallisation at 110° C. It is easily soluble in water, and the solution shows mutarotation. Its final specific rotatory power at 20° C. is $[\alpha]_D = +52{\cdot}74°$. It ferments with yeast.

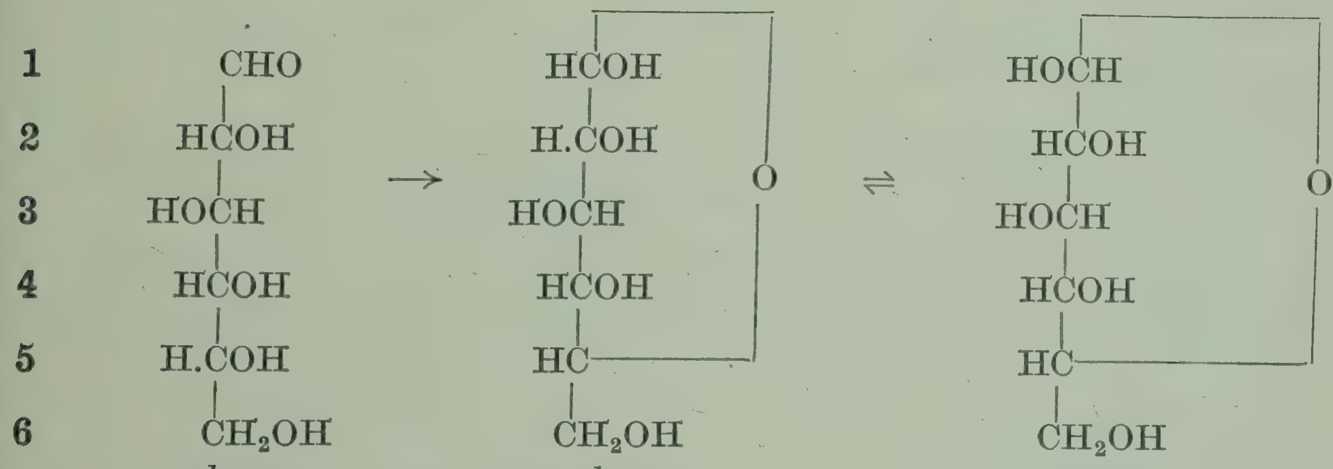
d-(−)-FRUCTOSE or LÆVULOSE occurs mixed with glucose in honey and in fruit sugar. It is also, with glucose, formed by the digestion or inversion of cane sugar. It is crystallisable with difficulty. Its watery solution is lævo-rotatory, and reduces Fehling's solution somewhat less strongly than glucose, its reducing power being 92, if we take that of glucose as 100. It ferments readily with yeast; with phenyl hydrazine it gives the same osazone as is formed from glucose.

d-(+)-GALACTOSE is formed by the hydrolysis of milk sugar or the galactosides of the brain with dilute mineral acids, and by the hydrolysis of certain vegetable gums. It is much less soluble in water than glucose. It is dextrorotatory and shows marked mutarotation. With ordinary yeast it ferments, but extremely slowly. Its reducing power is somewhat less than that of glucose.

MANNOSE. Mannose, though an assimilable sugar, is of such rare occurrence in our foodstuffs that it plays practically no part in animal physiology. It is dextro-rotatory, reduces Fehling's solution, ferments easily with ordinary yeast, and gives an osazone which is identical with that derived from glucose.

The straight-chain formulæ already given for the hexoses suffice to explain the majority of the reactions of the sugars; but a study of the behaviour of the reducing sugars in solution indicates that they possess a structure more labile than is suggested by the conventional static formula. A sample of glucose recrystallised from water exhibits when dissolved in water, a progressive decrease in rotatory power, during which the specific rotation falls from about 100° to a constant value of 52·7°. On the other hand, glucose recrystallised from pyridine gives, upon solution in water, an initial specific rotation of about 20°, rising slowly to 52·7°. This phenomenon receives the name of mutarotation, and is due to the fact that in solution glucose exists as an equilibrium mixture of two optical isomers which are designated α and β glucose respectively. α Glucose is the isomer usually handled as solid glucose. On the other hand, β glucose is obtained when the sugar is crystallised from pyridine.

Structural investigation of the two forms has shown that a lactone ring bridges the first and fifth carbon atoms, thus creating a new asymmetric carbon atom at position 1, and making it thus possible to have two isomers, the α and β forms, by altering the positions of the H and OH at position 1.



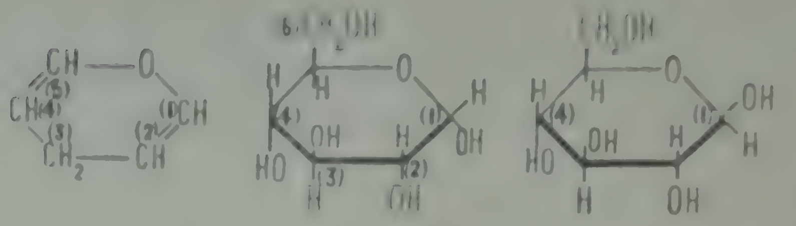
glucose (straight-chain formula)

α-glucose

β-glucose

Owing to the angle at which the bonds operate, the sugars are more correctly considered as derived from a six-atom heterocyclic ring called Pyran. According to the nomenclature proposed by Haworth these sugars are called pyranoses, *e.g.* glucose could exist as α or β gluco-pyranose; a 1 : 4 or γ lactone (butylene oxide) linkage can also exist, being a very reactive form called γ glucose (or gluco-furanose, owing to its relation to the 5-atom ring called Furan) of which again there are α and β forms. It is similar with other sugars.

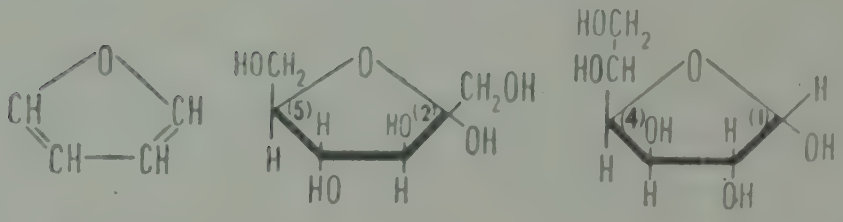
The formulæ, which are now conventional, given below, exhibit this ring structure for glucose and fructose: a carbon atom is to be assumed at each corner, except where the O appears. The diagrams suggest the spatial arrangement in relation to the carbons.



Pyrane

α-glucose (α-gluco-pyranose)

β-glucose (β-gluco-pyranose)



Furane

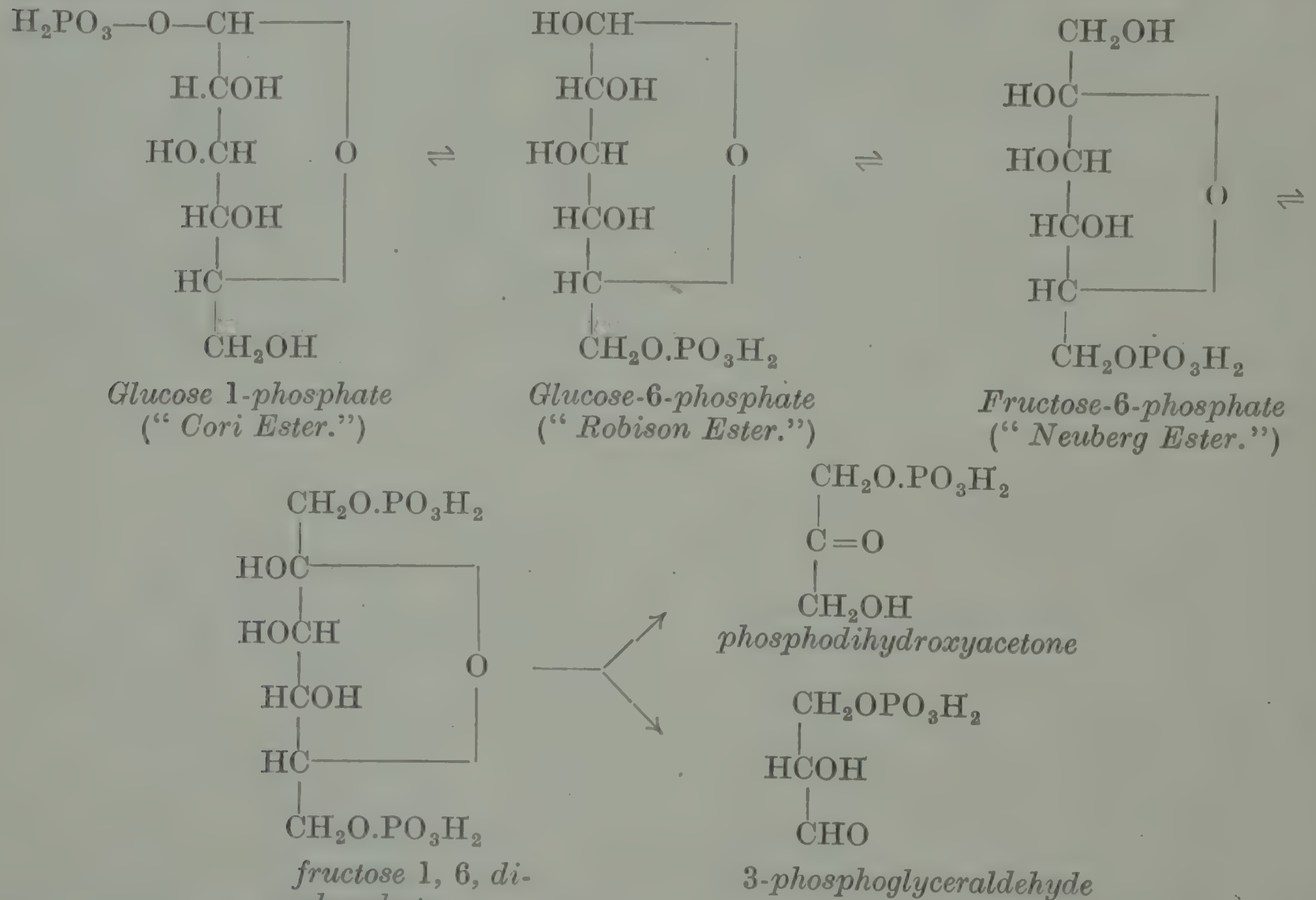
fructose (α-fructofuranose)

α-gluco-furanose

The figures shown in brackets are the conventional position-numbers.

DERIVATIVES OF THE HEXOSES

The phosphoric esters of the hexoses play an important part in the intermediary metabolism of carbohydrates, since they break down into triose phosphates (*v.* p. 31) :—



Glucose 1-phosphate ("*Cori Ester.*")

Glucose-6-phosphate ("*Robison Ester.*")

Fructose-6-phosphate ("*Neuberg Ester.*")

phosphodihydroxyacetone

fructose 1, 6, di-phosphate ("*Harden-Young Ester.*")

3-phosphoglyceraldehyde

The so-called "Embden ester" is a mixture of glucose-6- and fructose-6-phosphates.

Asymmetric syntheses of sugars from phosphorylated derivatives have been brought about by the use of enzymes from muscle and yeast ; these bring about an aldol condensation between dihydroxyacetone phosphate and an aldehyde.* Thus with *d*-glyceraldehyde, fructose 1-phosphate results. It is probably by some such means that sugars are interconverted in the body.

Two other derivatives of glucose are of physiological importance, namely, glucosamine and glucuronic acid.

Glucosamine has the structural formula :—

$$CH_2OH.(CHOH)_3.CHNH_2.CHO$$

---

* Lohmann. *Angew. Chem.*, 1936, **49**, 327.

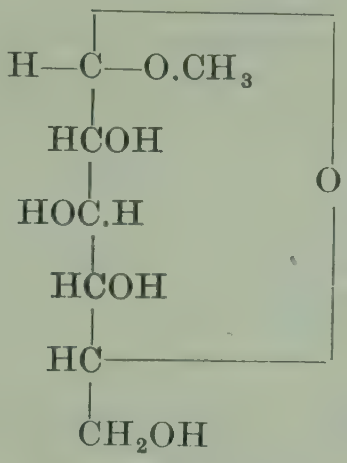
It is obtained from chitin, which forms the exoskeleton of large numbers of the invertebrata, by boiling this with concentrated hydrochloric acid. It is obtained as a decomposition product of the glucoproteins, such as the mucins. In solution it is dextro-rotatory, reduces Fehling's solution, and gives an osazone resembling that derived from glucose.

Glucuronic Acid may be regarded as one of the first results of oxidation of the glucose molecule. The group which has undergone oxidation is not the readily oxidisable CHO group, but the CH_2OH group at the other end of the molecule. The formula of this acid is therefore :

$$COOH.(CHOH)_4.CHO$$

It may be produced by the oxidation of glucose by means of hydrogen peroxide. In the free state it does not occur in the animal body. It is constantly found in the urine after administration of certain drugs such as phenol, camphor or chloral, and then occurs as a conjugated acid with these substances. These glucuronates are lævo-rotatory, though the free acid is dextro-rotatory. In the free state it reduces Fehling's solution and gives an osazone which is not sufficiently characteristic to distinguish from glucosazone. It does not undergo fermentation with yeast.

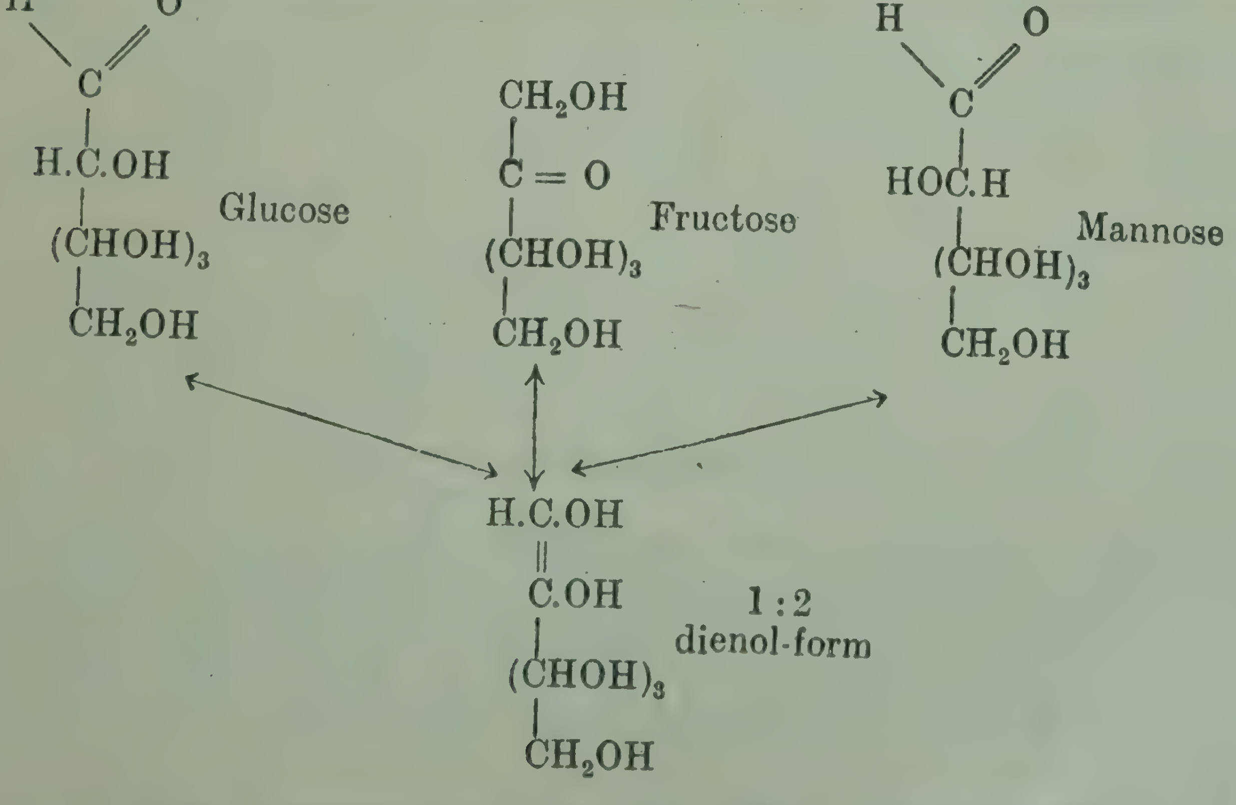
Glycosides. The glycosides are ethers formed by the reducing group of sugars. Thus by methylating α-glucose we get α-methyl glucoside (the glucosides being the glycosides derived from glucose, galactosides from galactose, etc.).



α-methylglucoside

A large number of glucosides occur as plant products—amygdalin, salicin, phlorrhizin indican, saponin—and appear to be β glucosides, *i.e.* are ethers formed from β glucose and an alcohol. The disaccharides are also glucosides : thus maltose is glucose α glucoside, whilst lactose is probably glucose β galactoside.

The Action of Alkali on the Hexoses. A trace of alkali greatly accelerates the mutarotation of the hexoses, indicating an increased lability of the carbon atom at position 1. In more strongly alkaline solutions further changes occur. In the case of glucose a progressive fall in rotatory power reaching a lævo-rotation will be observed. Examination of the solution will then show the presence not only of glucose but also of fructose and mannose.* Since the differences between these three hexoses are confined to the carbon atoms 1 and 2, it will be seen that their ready interconversion in alkaline solution may be explained as occurring through the intermediate production of a 1 : 2 dienol-form.

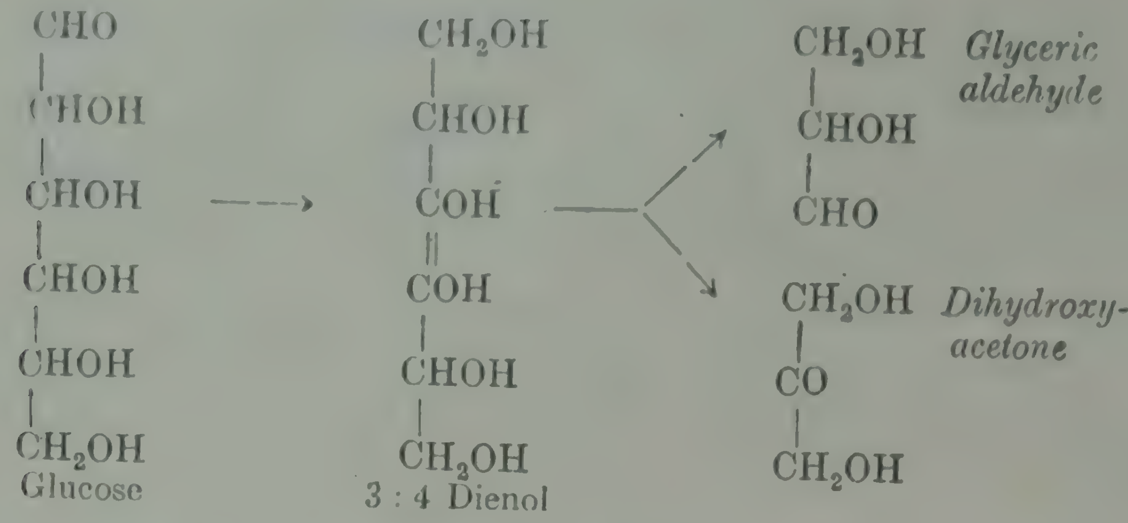


---

* Nef. *Annalen der Chemie*, 1910, **376**, 1 ; 1914, **403**, 204.

It is of interest that these three hexoses, which are most nearly related *in vitro*, are als the most readily assimilated by the animal. The difference is that, whereas alkali effect an indiscriminate interconversion of glucose, fructose and mannose, the living cell show directional activity in turning them all into glucose. The anomalous position of galactos as an assimilable sugar is not capable of so facile an explanation on structural lines.

In somewhat stronger alkaline solutions and in absence of oxygen, the 3 : 4 dienol-form results, and this readily breaks up into two important 3-carbon derivatives, glyceric aldehyd and dihydroxyacetone :—



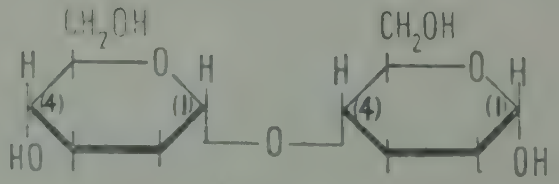
In strongly alkaline solutions in presence of oxygen still more extensive changes occur. A large number of different acids containing varying numbers of carbon atoms, such as oxalic, lactic, tartaric and saccharic acids, and complex resinous substances are formed. There is so much in common between the alkaline and the biochemical degradation of carbohydrate that one could expect that the same broad lines are followed in each, though in the living cell the notable distinction is the production of a selected few of the substances appearing in alkaline solution *in vitro.*

THE DISACCHARIDES ($C_{12}H_{22}O_{11}$)

The disaccharides are formed by the union of two molecules of monosaccharides with the elimination of one molecule of water, and can be regarded as glucosides, galactosides, etc. On hydrolysis, *e.g.* on heating with acids, they take up one molecule of water and are split up into the corresponding monosaccharides. Thus cane sugar gives glucose and fructose, maltose gives two molecules of glucose, while lactose gives glucose and galactose.

Cane Sugar, or sucrose, is widely distributed throughout the vegetable kingdom, and forms an important article of diet. It has no reducing power, is strongly dextro-rotatory and has a specific rotatory power of + 66·5°. On hydrolysis it is converted into glucose and fructose. Owing to the fact that fructose rotates polarised light more strongly to the left than glucose does to the right, the mixture of the two monosaccharides so obtained is lævo-rotatory. On this account the change to the mixture of monosaccharides is known as *inversion*. With yeast, cane sugar first undergoes inversion by a special enzyme present in the yeast (*invertase*), and the mixture of fructose and glucose is then fermented. Invertase is also present in the intestinal juice.

Maltose is formed during the hydrolysis of starch by acids or by the enzyme amylase, and is the chief sugar in germinating barley or malt. It is α-glucose-α-glucoside :—



Maltose
(abbreviated conventional formula)

It is strongly dextro-rotatory, ferments easily with yeast, and reduces Fehling's solution ; its reducing power is about 70 per cent. of that of glucose. With phenyl hydrazine it gives phenyl maltosazone, which forms characteristic yellow crystals with a melting-point of 206° C. The enzyme *maltase*, of the intestinal juice, converts it to glucose.

Lactose is a constituent of milk. It is much less soluble in water and much less sweet than are the two other disaccharides. It is dextro-rotatory, shows mutarotation and is not fermented by ordinary yeast. Lactose can be split by the agency of an enzyme, *lactase*, which occurs in the animal body and in certain moulds, into glucose and galactose. Lactose reduces Fehling's solution and gives an osazone, which is easily soluble in hot water.

THE POLYSACCHARIDES $(C_6H_{10}O_5)n$

These are complex substances of high molecular weight formed by combination of monosaccharide molecules, with elimination of water. In the animal body the sole significance of polysaccharides lies in their value as foodstuffs. In plants, anhydrides both of hexoses (hexosans $(C_6H_{10}O_5)_n$) and pentoses (pentosans $(C_5H_8O_4)_n$) occur in bewildering variety. Unlike the other carbohydrates, the higher polysaccharides are not destroyed by hot strong alkaline solutions. The chemical constitution of the polysaccharides is a difficult problem because of their complexity and their colloidal behaviour. The methylation methods of Irvine have given us a deeper insight into the manner of combination of the hexoses in a polysaccharide molecule. It appears certain that glycogen and starch are both composed of α-glucopyranose molecules, linked up as maltose units.

The enzymic synthesis of polysaccharides, amylose, starch and glycogen has been effected, the starting-point being glucose 1-phosphate.*

Starch is present in large quantities in nearly all vegetable foods, and is an important constituent of the cereals, as well as of tubers, such as the potato. In the plant cells it occurs as concentrically striated grains, insoluble in water. In hot water the grains swell up and burst, forming a thick paste, which sets to a jelly on cooling. This colloidal solution, as well as the original starch grains, gives an intense blue colour on the addition of iodine. On treating with cold dilute acid, it is converted into a soluble modification, the so-called soluble starch, which also gives a blue colour with iodine. This modification is likewise produced as the first stage of the action of the enzyme *amylase* upon starch. On boiling with dilute acids, starch is converted first into a mixture of dextrins, then into maltose, and finally into glucose.

Natural starch actually consists of two polysaccharides: 25 per cent. of it is *amylose*, giving a blue with iodine, and 75 per cent. is *amylopectin*, which turns purple with iodine. They occur in the ratio 2 : 1, of molecular weights about 60,000 and 300,000 respectively. Each consists of about 24 glucose units, the high molecular weights being due to aggregation. Amylopectin is closely associated with phosphoric acid, with which it may indeed be in combination. On acid hydrolysis both yield glucose; on hydrolysis with barley diastase, amylose yields maltose, but amylopectin is only partially depolymerised, though oat diastase can split it also finally into maltose. On acting upon starch with amylase, such as the diastase of germinating barley, or with the amylase occurring in saliva or pancreatic juice, it undergoes hydrolysis, the final result of the action being a mixture of four parts of maltose to one part of stable dextrin, the latter being derived from the amylopectin. In the intermediate stages in this reaction the first product is soluble starch or *amylodextrin*. This breaks up into a reducing sugar and another dextrin, *erythrodextrin*, which gives a red colour with iodine; and this dextrin, on further hydrolysis, yields reducing sugar and *achroödextrin*, which is not coloured by the addition of iodine. Thus there is a series of successive hydrolytic decompositions of the molecule, each resulting in the splitting off of a molecule of sugar and the production of a lower dextrin.

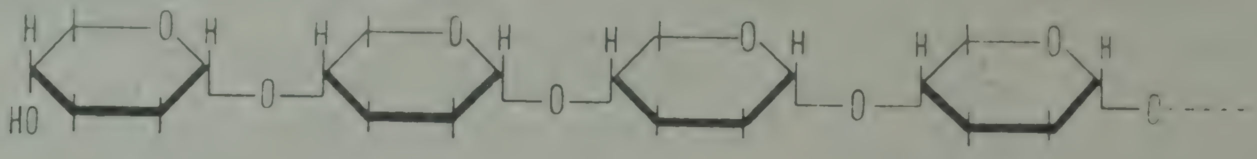
The **Dextrins** are ill-defined bodies which are difficult to separate. They are amorphous white powders, easily soluble in water, forming gummy solutions when concentrated. They are insoluble in alcohol and ether. With cupric hydrate and caustic alkali they form blue solutions, which reduce slightly on boiling. They are not precipitated by saturation with ammonium sulphate. On boiling with dilute acids, they are converted entirely into glucose.

---

* Cori *et al.* *Science,* 1939, **89,** 464. Hanes. *Proc. R. S.,* 1940, **128B**, 421; **129B**, 174.

INULIN, which somewhat resembles starch, occurs in dahlia tubers. It is easily hydrolysed by weak acids into d-fructose.

GLYCOGEN is a polysaccharide which closely resembles starch in structure and properties. It is found in the liver, muscles and other tissues of the body, in all fœtal tissues, and in yeast. It is a white powder forming an opalescent solution in water, and is precipitated from its solution on the addition of alcohol to 60 per cent., by strong acetic acid, by basic lead acetate, or by saturation with solid ammonium sulphate. On boiling with acids, it is entirely converted into glucose. It is affected by amylase in the same way as starch, giving first dextrins and finally maltose. With iodine it gives a mahogany-red colour which, like the blue colour produced in starch, is destroyed by boiling, to return again on cooling. There are two forms of glycogen, one with 12 and one with 18 α-glucopyranose units.



Scheme of portion of glycogen molecule comprising 4 α - glucose units.

THE CELLULOSES. Cellulose is a colourless, insoluble material, or mixture of materials, which forms the cell walls of plants, and is therefore a constituent of most of our vegetable foods. It is insoluble in water or dilute acids or alkalies, but soluble in ammoniacal cupric oxide solution. On boiling with strong acids, it gradually undergoes hydrolysis and yields glucose. In herbivorous animals cellulose is digested by the agency of micro-organisms which swarm in the paunch of ruminants and in the cæcum of other herbivora. By these the cellulose is converted into lower fatty acids, *e.g.* acetic and propionic, which are changed in the body into fats and sugar. In man cellulose undergoes practically no change in digestion, and serves merely by its bulk to promote peristalsis and the normal evacuation of the bowels.

For Reference

ARMSTRONG, E. F., and ARMSTRONG, K. F. (1931). The Glycosides. (Longmans, Green & Co.)

ARMSTRONG and ARMSTRONG (1934). The Carbohydrates. (Longmans, Green & Co.)

BELL (1948). Introduction to Carbohydrate Biochemistry. (London. Univ. Tutorial Press.)

HAWORTH (1929). The Constitution of the Sugars. (Edw. Arnold & Son.)

3. THE PROTEINS

The proteins form the most important constituent of protoplasm, and must always be present in a diet. Although characterised by certain chemical tests, it is not easy to give a simple definition in chemical terms of what we mean by the term "protein."

Elementary Composition. All proteins contain oxygen, hydrogen, nitrogen (15–17·6 per cent.), carbon and, usually, sulphur (0·3–2·2 per cent.).

Physical Characters. The proteins are tasteless colloids, many of which can be crystallised. Most of them are soluble either in water, weak salt solutions, or in dilute acids or alkalies. They tend to become insoluble at the iso-electric point, and this change is sometimes found to be irreversible. They are extremely liable to undergo the change called *denaturation*, the nature of which is obscure, as a result of many kinds of chemical or physical treatment. This denaturation may result from shaking, change of temperature, alteration of reaction, or addition of neutral salts. *Coagulation* by heat forms a distinguishing feature of a number of members of this class, which are therefore spoken of as "coagulable proteins." For instance, white of egg is a solution of different proteins. If a strong solution be boiled the whole fluid sets to a solid white mass or *gel*. This change is irreversible, *i.e.* it is not possible by lowering the temperature to bring the white of egg again into solution, and many properties of the protein have been changed in the act of coagulation. A few proteins, *e.g.* gelatin, form gels at lower temperatures which melt reversibly on warming.

THE STRUCTURE OF THE PROTEIN MOLECULE

We can arrive at some idea of the manner in which the protein molecule is built up by breaking it down by hydrolysis.

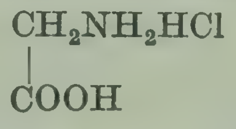
By prolonged acid hydrolysis we obtain a complex mixture of substances, nearly all of which are α-amino-acids, and must be regarded as the proximate constituents of the protein molecule.

A similar hydrolytic change may be effected by the use of digestive enzymes, pepsin, trypsin, erepsin, and so on. These enzymes are all milder in their action than the strong acids. Pepsin for instance effects only a partial decomposition of the protein molecule. Its action results in the formation of substances which still present all the protein reactions and are classified as hydrated proteins or as proteoses and peptones. Trypsin carries the protein a stage further and gives a mixture of polypeptides and amino-acids. Erepsin breaks the polypeptides down to amino-acids.

THE AMINO-ACIDS

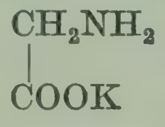
Nearly all the known disintegration products of the main protein molecule are amino-acids. They are soluble in water and easily diffusible. All the amino-acids, except glycine, derived from proteins are optically active, whereas those obtained by synthesis are inactive, or racemic, mixtures of the *d*- and *l*-forms.

The presence in the amino-acids of the basic radical NH_2 and of the acid group COOH lends to these bodies a double character. The monoamino-monocarboxylic acids, possessing neither acid nor alkaline reaction, are able in the presence of strong acids or bases to act either as base or acid. Thus, with hydrochloric acid, glycine forms glycine hydrochloride:



$$\begin{array}{l} CH_2NH_2HCl \\ | \\ COOH \end{array}$$

With bases these amino-acids form salt-like compounds such as potassium amino-acetate



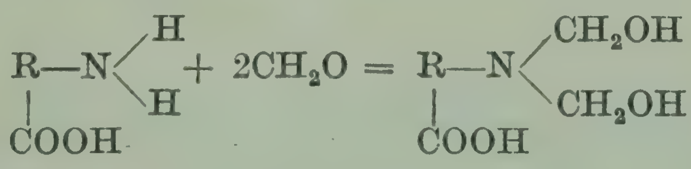
$$\begin{array}{l} CH_2NH_2 \\ | \\ COOK \end{array}$$

On account of the dual nature of the amino-acid molecule, these substances act as feeble electrolytes. Since they may act either as cations or anions, they are spoken of as *amphoteric* electrolytes (*v.* also p. 76).

Amino-acids also combine as acids and bases with one another. This power of combination much increases the difficulty of separating the constituents from a mixture of amino-acids. Amino-acids, which singly are sparingly soluble, are often soluble when in the presence of other amino-acids.

The diamino-acids, having two amino groups, are definitely basic in nature, while the dicarboxylic acids, with two carboxyl groups, are acids.

When an amino-acid is treated with formaldehyde the amino group is attacked, and the carboxyl group then has full acidic power and can be titrated.



$$\begin{array}{l} R\text{—}N\begin{array}{l} \diagup H \\ \diagdown H \end{array} \\ | \\ COOH \end{array} + 2CH_2O = \begin{array}{l} R\text{—}N\begin{array}{l} \diagup CH_2OH \\ \diagdown CH_2OH \end{array} \\ | \\ COOH \end{array}$$

Since each COOH group liberated corresponds to one NH_2 group fixed, this method is often used for estimating the amino-groups in any mixture.

An important property of amino-acids is that of combining with carbon dioxide to form carbamino-acids, *e.g.*:—

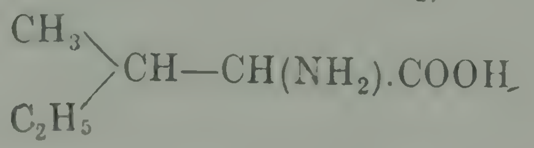
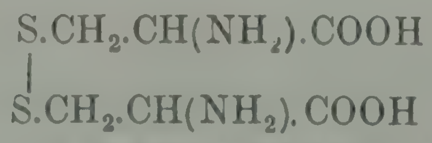
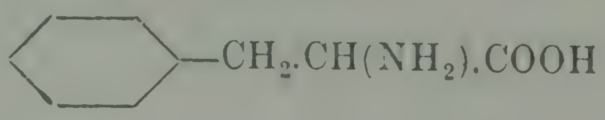
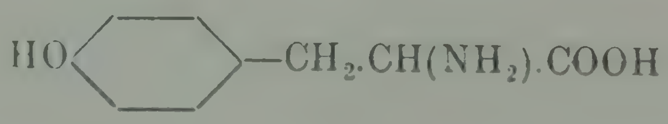
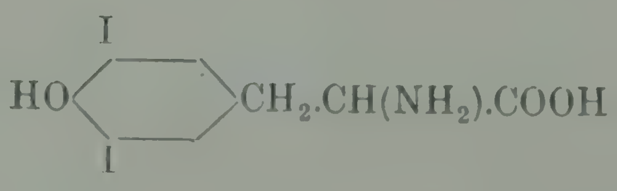
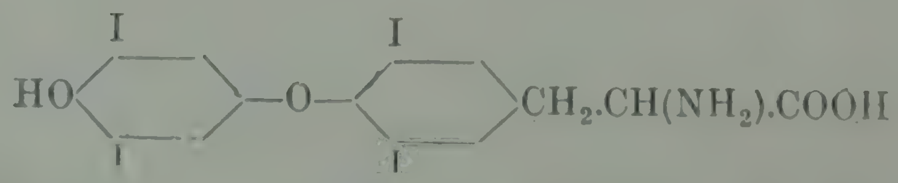
$$R.NH_2 + CO_2 = R.NH.COOH,$$

and this property is shared by certain proteins, which are thereby enabled to combine with carbon dioxide.

Since proteins are built up from a large number of amino-acids, it is not surprising to find that they likewise are amphoteric substances. The predominance of acid or basic properties will depend on whether the protein contains a surplus of dicarboxylic acids or of di-amino acids. It follows that the proteins can exist as salts with either acids or bases or in the free state. The factor which determines the condition of the protein is the hydrogen ion concentration. It has been found that the physical properties of protein solutions vary greatly with the hydrogen ion concentration and exhibit abrupt maxima or minima at a definite reaction which is characteristic of the protein. It seems probable that at this point —the *iso-electric point*—the protein is combined neither with acids nor with bases, but exists as a very feebly ionised free protein. On either side of the iso-electric point a highly ionised protein salt of some sort will exist, the nature of which again depends on the hydrogen ion concentration. These characteristic changes of the protein molecule with the reaction of the medium are of considerable importance.

The α-amino-acids which have been isolated from the products of hydrolysis of proteins are so varied that it may be a convenience to append a list of the physiologically important amino-acids which have actually been separated by hydrolysis of proteins. (The commonly used names are given on the left.) *

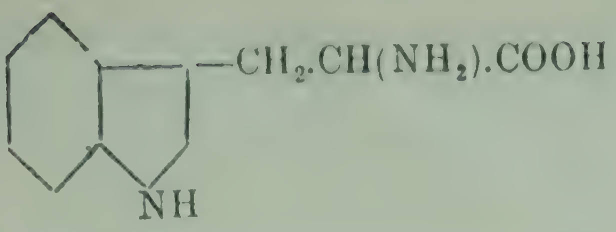
Monoamino monocarboxylic (or "Neutral") Amino-Acids.

Glycine.	$NH_2.CH_2COOH$	α-Amino acetic acid
d-Alanine.	$CH_3.CH(NH_2).COOH$	α-Amino propionic acid
l-Serine.	$CH_2(OH).CH(NH_2)COOH$	β-Hydroxy-alanine
l-Threonine.	$CH_3.CH(OH).CH(NH_2).COOH$	β-hydroxy-α-amino butyric acid
d-Norvaline.	$CH_3.CH_2.CH_2.CH(NH_2)COOH$	α-amino valeric acid
d-Valine.	$(CH_3)_2CH.CH(NH_2).COOH$	ββ-Dimethyl alanine
d-Norleucine.	$CH_3.CH_2.CH_2.CH_2.CH(NH_2).COOH$	α-amino caproic acid
l-Leucine.	$(CH_3)_2CH.CH_2.CH(NH_2).COOH$	β-Isopropyl alanine
d-Isoleucine.	 $\genfrac{}{}{0pt}{}{CH_3}{C_2H_5}{>}CH-CH(NH_2).COOH$	ββ-Methyl ethyl alanine
l-Cystëine.	$HS.CH_2.CH(NH_2).COOH$	β-Thio-alanine
l-Cystine.	 $\begin{array}{l}S.CH_2.CH(NH_2).COOH\\ \vert\\ S.CH_2.CH(NH_2).COOH\end{array}$	Dicystëine
l-Methionine.	$CH_3.S.(CH_2)_2CH(NH_2)COOH$	γ-Methylthiol — α-Amino butyric acid
l-Phenyl alanine.		Phenyl alanine
l-Tyrosine.		ρ-Hydroxyphenyl alanine
3, 5, Diiodotyrosine.		
Thyroxine.		

---

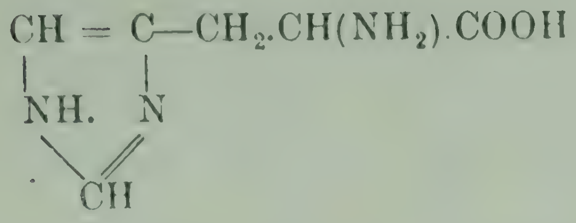
* For nomenclature *v.* *Biochem. J.*, 1948, **42**, 1. The prefix *d*- or *l*- here refers to the actual rotation. Confusion has sometimes been caused by the use of the Fischer configurative designation, as described for carbohydrates (p. 32). According to this, most of the natural amino-acids belong to the l- series: the actual rotation is then indicated by (+) or (−) following the letter. Thus, natural glutamic acid would be *l* (+), and serine *l* (−), etc., but natural threonine would be *d* (−).

l-Tryptophan.



β-Indol alanine *

l-Histidine.



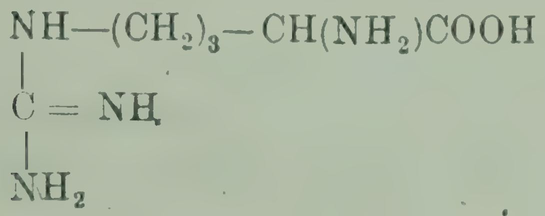
β-Iminazol alanine

Diamino- (or Basic) Amino-Acids

d-Ornithine. $NH_2(CH_2)_3CH(NH_2).COOH$ αδ-Diamino-valeric acid

Citrulline. $NH_2.CO.NH.(CH_2)_3.CH(NH_2).COOH$

d-Arginine.



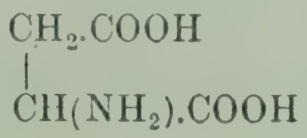
$$\begin{array}{l} NH{-}(CH_2)_3{-}CH(NH_2)COOH \\ | \\ C = NH \\ | \\ NH_2 \end{array}$$

α-amino-δ-guanidino-valeric acid

d-Lysine. $NH_2(CH_2)_4CH(NH_2).COOH$ α-Diaminocaproic acid

Dicarboxylic (or Acidic) Amino-Acids

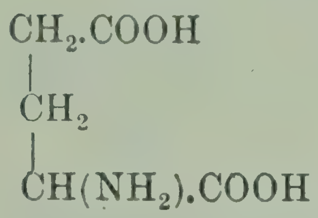
l-Aspartic Acid.



$$\begin{array}{l} CH_2.COOH \\ | \\ CH(NH_2).COOH \end{array}$$

Aminosuccinic acid

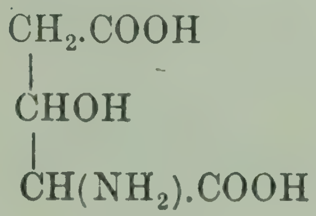
d-Glutamic Acid.



$$\begin{array}{l} CH_2.COOH \\ | \\ CH_2 \\ | \\ CH(NH_2).COOH \end{array}$$

α-Aminoglutaric acid

d-Hydroxyglutamic Acid.

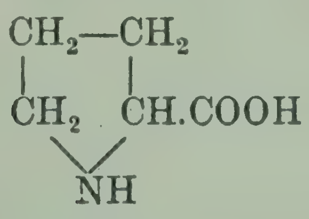


$$\begin{array}{l} CH_2.COOH \\ | \\ CHOH \\ | \\ CH(NH_2).COOH \end{array}$$

α-Amino, β-hydroxyglutaric acid

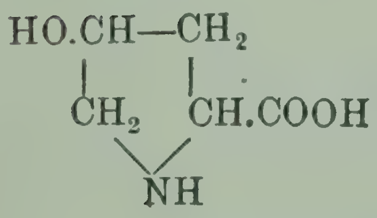
"Condensed" Amino-Acids

l-Proline.



α-pyrrolidine carboxylic acid.

l-Hydroxyproline.



γ-hydroxy-α-pyrrolidine carboxylic acid

Several of the chemical tests for proteins depend on the presence of specific amino-acids; for example, the Millon's reaction on the presence of tyrosine, the glyoxylic test on the presence of tryptophan, etc.

Under the action of bacteria, amino-acids may undergo a process of deamination which may be accompanied by simple hydrolysis or by reduction. In the former case an amino-acid may be converted into an hydroxyacid, in the latter case into a fatty acid.

---

* Synthesis. *v.* Snyder and Smith. *J. Amer. Chem. Soc.*, 1944, **66**, 350.

On the other hand, some bacteria may split off carbon dioxide from the amino-acids and yield amines. Thus tyrosine becomes *p*-hydroxyphenylethylamine, and histidine becomes *histamine*, a substance having important physiological effects (p. 97).

OTHER CONSTITUENTS OF THE PROTEIN MOLECULE

When we add together the total amino-acids obtainable by the acid disintegration of any given protein, a considerable proportion of the original protein is often unaccounted for. Much of this remainder may represent analytical loss, but some of it may consist of substances not as yet identified. One substance sometimes present is glucosamine. Proteins belonging to the class of mucoids yield on acid hydrolysis from 16 to 30 per cent. glucosamine. Egg albumin, serum albumin and serum globulin all contain in their molecules a complex carbohydrate consisting of glucosamine and mannose.*

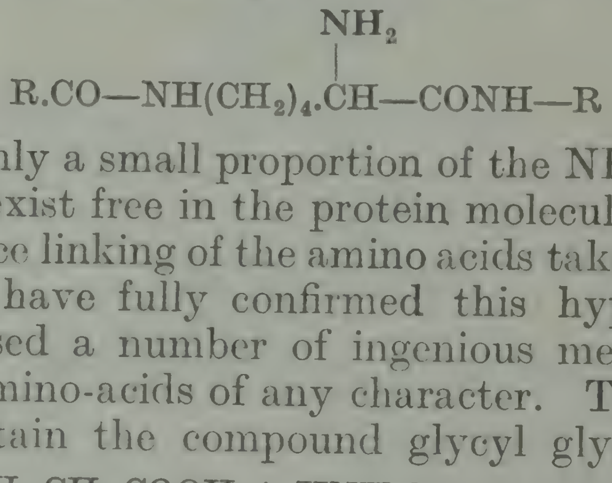
THE BUILDING UP OF THE PROTEIN MOLECULE

The amino-acids are present in different proportions in the individual proteins, so that in many cases a large number of identical amino-acid groups must be present in the protein molecule, with smaller numbers of other groups. One method of linkage of amino-acids to form larger molecules is by condensation of a COOH of one amino-acid and an NH_2 group of another to form the *peptide* linkage :

$$—CO.NH—$$

It is generally believed that, in this manner, molecule can be added to molecule until finally the gigantic molecule of a protein results. In such molecules there is believed to be a regular pattern of amino-acids, each of which recurs with constant frequency at a particular place in the chain.†

Of course, each lysine group in a peptide will give one free NH_2 group :—



$$R.CO—NH(CH_2)_4.\overset{\overset{NH_2}{|}}{C}H—CONH—R$$

but apart from these, only a small proportion of the NH_2 groups present in the separated amino-acids exist free in the protein molecule, and these must be at the end of the chain, since linking of the amino acids takes place by NH groups. Synthetic experiments have fully confirmed this hypothesis. Fischer, and later Bergmann,‡ devised a number of ingenious methods for combining together two or several amino-acids of any character. Thus from two molecules of glycine we may obtain the compound glycyl glycine, as follows :

$$NH_2.CH_2.COOH + HNH.CH_2.COOH =$$
$$NH_2.CH_2.CO.NH.CH_2.COOH + H_2O$$

These compounds were designated *peptides* by Fischer. He distinguished di-, tri-, tetra, etc., or *polypeptides* according to the number of individual amino-acids taking part in the formation of the compound, *e.g.* alanyl leucine, glycyl tyrosine, dialanyl cystine, dileucyl cystine, leucyl pentaglycyl glycine, and so on. The last named would be built up out of one molecule of leucine, and six molecules of glycine. Abderhalden built up a molecule containing eighteen constituent amino-acids. These polypeptides are still very far below the proteins in molecular size, and proteins have not been made artificially. The higher polypeptides closely resemble the peptones. Most of them, even if derived from relatively insoluble amino-acids, are soluble in water, insoluble in alcohol. They dissolve in mineral acids and in alkalies, with the formation of salts, and have a bitter taste. The higher members of the series give certain reactions, such as the biuret reaction, which are regarded as characteristic of peptones, and like the latter are precipitated by phosphotungstic acid. Some are split up by certain proteinases,* and by erepsin, into amino-acids.

---

* RIMINGTON. *Biochem. J.*, 1929, **23**, 430.

† BERGMANN and NIEMANN. *J. Biol. Chem.*, 1937, **118**, 301 ; 1938, **122**, 577.

‡ BERGMANN. *Science*, 1934, **79**, 439. BERGMANN and ZERVAS. *J. Biol. Chem.*, 1936, **113**, 341.

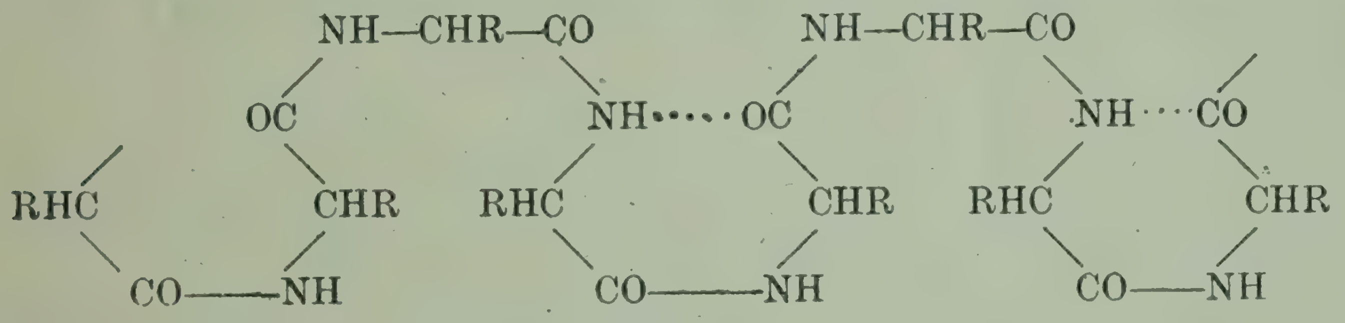
Thus in the building up of the protein molecule there is an almost unlimited capacity for coupling up of numerous amino-acid groups, the link in each case being the nitrogen. Of the two or more possible optical isomers of each amino-acid containing more than two carbon atoms, only one is found in nature.

A still further flexibility in its reactions is conferred on the protein molecule by intramolecular rearrangement occurring with great readiness in its component molecules. As a result of this the protein molecule may not be a single peptide chain, but a system of several chains, differing somewhat in composition, and linked to one another, not by amino groups, but by carboxyl, or imide, or thiol-ester linkages. Thus, according to Chibnall, egg-albumin may consist of at least 4 peptide chains, having free terminal amino-groups (*i.e.* not accountable for from the lysine content), and lactoglobulin may have 9 peptide chains, so explaining the presence of 8 extra free amino-groups, which must presumably be terminal in position.

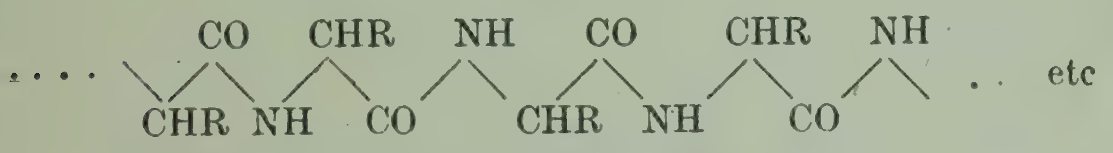
The fact that pepsin will act on proteins, but not on polypeptides, is a further indication that the peptide structure may not be the only type of linkage by which the constituent amino-acids are assembled in the protein molecule. If we consider that perhaps some hundred of the amino-acid groups may go to making up a single protein molecule, it is possible to form some conception of the enormous variability in structure possible to such a compound.

The study of the ultrastructure of proteins by the use of X-ray spectra and by other physical methods, such as the study of thin films, orientation in electric fields, etc., gives another viewpoint from which to study the architecture of their molecules. These investigations show two types of molecular structure, *fibrous* and *globular*.

Among the fibrous proteins are those of myosin, keratin, and fibroin, present respectively in muscle, hair and silk. In these we have essentially a long polypeptide chain. In the normal state the chain is kinked owing to electrostatic attraction between side-chains, *e.g.* in α-keratin :—



α-keratin



β-keratin (stretched)

but when stretched the chain becomes more regular (β-keratin), and if heated in the stretched state this form becomes denatured and permanent (" permanent waving "). The myosin of muscle is probably similar.

In the globular protein molecules, the kinking is exaggerated, so that the molecule folds on itself in a regular pattern. Such proteins are insulin, pepsin, lactoglobulin and hæmoglobin. According to Wrinch,† regular geometrically constructed molecules may result (cyclol structure), containing 6, 18, 30, 42, 54, 66, 72, 90 or more amino-acids, which ultimately form a globular structure. In these molecules it is assumed that some of the peptide linkages are tautomerised to — C(OH) : N — linkages.

THE CONSTITUTION OF DIFFERENT PROTEINS

Nearly all the proximate constituents of the main protein molecule, so far as we know, are amino-acids. Of these the following have been isolated, namely, glycine, alanine, valine, leucine, isoleucine, proline, hydroxyproline, serine, threonine, phenyl alanine, glutamic acid, β-hydroxyglutamic acid, aspartic acid, tyrosine, tryptophan, cystine, methionine, lysine, histidine and arginine.

---

* Bergmann, Fruton and Fraenkel-Conrat. *J. Biol. Chem.*, 1937, **119**, 35.

† Wrinch. *Proc. R. S.*, 1937 (A), **160**, 63 ; **161**, 505.

The question now arises whether all the varieties of protein owe their peculiarities to the presence of different amino-acids, or whether the greater number of the amino-acids above mentioned are present in all proteins, the differences between the latter being determined by differences in the arrangement and relative amounts of these proximate constituents. A large number of analyses of various proteins have been made by Osborne, Dakin, Chibnall and others, utilising the methods for the isolation of amino-acids devised by Fischer and by Dakin. The constitution of some representative proteins as determined in this way is given in the Table below.

	Serum albumin.	Egg albumin.	Edestin (hemp seed).	Gliadin.	Caseinogen.	Globin.	Salmin.	Sturin.	Gelatin.	Keratin (from horse hair).	Zein.
Glycine .	0	0	3·8	0·9	0·45	0	0	—	25·5	4·7	0·0
Alanine .	2·7	8·1	3·6	2·7	1·85	4·2	—	—	8·7	1·5	9·8
Serine .	0·6	—	0·33	0·12	0·5	0·6	7·8	—	0·4	0·6	1·0
Valine. .	—	—	present	0·3	7·93	—	4·3	—	—	0·9	1·9
Leucine .	20·0	7·1	20·9	6·0	9·7	29·0	0	—	7·1	7·1	25·0
Proline .	1·0	2·25	1·7	2·4	7·63	2·3	11·0	—	9·5	3·4	9·0
Hydroxyproline . .	—	—	2·0	—	0·23	1·0	—	—	14·1	—	—
Glutamic acid	7·7	16·1	20·7	46·92	21·77	1·7	—	—	5·8	3·7	31·3
Aspartic acid.	3·1	8·1	12·0	1·3	6·68	4·4	—	—	3·4	0·3	1·8
Phenyl alanine	3·1	4·4	2·4	2·6	3·88	4·2	—	—	1·4	0	7·6
Tyrosine .	2·1	4·1	4·3	2·4	4·5	1·5	—	—	0·01	3·2	5·7
Tryptophan .	present	1·3	1·5	1·0	1·50	present	—	—	—	—	0·0
Cystine .	2·3	1·8	1·4	0·25	?	0·3	—	—	—	10+	0·8
Lysine. .	—	5·1	2·4	0	7·62	4·3	0	12·0	5·9	1·1	0·0
Arginine .	—	5·6	16·7	3·4	3·81	5·4	87·4	58·2	8·2	4·5	1·8
Histidine .	—	1·5	2·4	1·7	2·5	11·0	0	12·9	0·9	0·6	0·8
β-hydroxy-glutamic acid.	—	—	—	—	10·0	—	—	—	0·0	—	2·5
Methionine .	—	5·1	2·4	—	—	—	—	—	—	—	—

These results show that all the proteins contain a very considerable proportion of the total number of amino-acids which have as yet been isolated from acid digests of proteins. The differences in various proteins cannot therefore be determined by qualitative differences in their constituent molecules, but must depend on the relative amounts of the amino-acids which are present and on their arrangement in the whole molecule. As regards relative amounts of amino-acids we find very striking variations. Thus glutamic acid, which forms 16 per cent. of egg albumin and only 1·7 per cent. of globin (derived from hæmoglobin), amounts to 36·5 per cent. in gliadin, the protein extracted from wheat flour. Striking differences are also noticeable in the relative proportions of the diamino-acids, the so-called hexone bases. Whereas in casein they form about 12 per cent. of the total molecule, in globin they form about 20 per cent. ; and in the protamines, salmin and sturin, about 85 per cent. of the total molecule consists of these bodies. On this account the two last-named proteins have a strongly basic character. From these figures it is evident also that certain of the amino-acids must occur many times over in the protein molecule. Thus in globin, if we assume the presence of one tyrosine molecule, there must be at least thirty-two leucine and ten histidine molecules. On these data the minimal molecular weight of hæmoglobin would come out at about 17,000.

THE MOLECULAR WEIGHT OF PROTEINS.—We may arrive at an approximate idea of the minimum size of the protein molecule in various ways, though all our calculations are apt to be vitiated by the difficulty of obtaining a preparation which is chemically pure, and by the ease with which molecules of the size which we must assume for proteins aggregate together or form adsorption combinations in varying proportions with other substances. Some idea of the molecular complexity represented by these weights may be gained by writing out the possible empirical formulæ of the various proteins, *e.g.* :

Egg albumin	$C_{204}H_{322}N_{52}O_{66}S_2$
Hæmoglobin	$C_{712}H_{1130}O_{245}N_{194}S_2Fe$

On the simplest empirical formula, therefore, most of the proteins have molecular weights of several thousand, *e.g.* hæmoglobin is at least 17,000. The application of physical methods, such as the rate of sedimentation at high-speed centrifugalisation, diffusion, viscosity, etc., shows that under most conditions these molecules are aggregated into larger complexes, which, however, can be broken up by altering the conditions. Most of these composite molecules have a molecular weight, according to Svedberg, which is a multiple of 17,600, as *e.g.* :—

Protein.	Molecular weight by sedimentation Velocity.		M(calc.)
Lactalbumin	17,500	...	17,600
Lactoglobulin	41,800	...	2 × 17,600
Insulin	40,900	...	2 × 17,600
CO-Hæmoglobin . . .	69,000	...	4 × 17,600
Serum globulin . . .	167,000	...	8 × 17,600
Hæmocyanin (*Helix*) . . .	6,630,000	...	384 × 17,600

(*v.* also p. 72).

THE CLASSIFICATION OF PROTEINS

At the present time it is impossible to make any classification of proteins based on their constitution, and we have therefore to use an arbitrary classification, based chiefly on the solubilities of the various proteins. Many proteins formerly considered to be chemical individuals have now been separated into various fractions. We shall here only indicate the characters of the main groups.

(1) **The Protamines.** These occur only in combination. They are obtained from the ripe spermatozoa of certain fishes, where they are in combination with nucleic acid. They are of relatively low molecular weight, *e.g.* 2,000, and are characterised by the very large amount of diamino-acids in their molecule. A small proportion of mono-amino-acids may also be obtained from their disintegration (*v.* Salmin, Table, p. 44). They possess strongly basic characters and form well-marked salts, *e.g.* sulphates, as well as double salts with platinum chloride. They contain no sulphur and do not coagulate on heating.

(2) **Histones.** This class of proteins, like the protamines, only occurs in combination with other substances, such for instance as nuclein and hæmatin. They may be obtained from red blood corpuscles, where they form the globin part of the hæmoglobin molecule, or from the thymus gland. The histones are precipitated from their watery solutions by addition of ammonia, but are soluble in excess of this reagent. In the presence of salts they are coagulated on boiling. The most characteristic feature is the high proportion of diamino-acids contained in their molecule, in consequence of which they are strongly basic.

(3) **Albumins.** These are soluble in water and are precipitated by complete saturation with ammonium sulphate or zinc sulphate.

Egg Albumin forms the greater part of the white of egg. It gives the ordinary protein tests, coagulates on heating at about 75° C., and is precipitated from its solutions if shaken with a drop of dilute acetic acid in excess of ether. It is lævo-rotatory, its specific rotatory power being — 35·5°.

Serum Albumins occur in large quantities in the blood plasma, serum, lymph, and tissue fluids of the body. They coagulate at 75° C., and are distinguished from egg albumin by greater specific rotatory power, – 56°, and by the fact that they are not precipitated by ether and acetic acid.

(4) **Globulins.** These bodies are insoluble in water and require the presence of a certain amount of neutral salt to dissolve them. They are precipitated from their solutions by complete saturation with magnesium sulphate or by half-saturation with ammonium sulphate. Important members of this class are:

Crystallin, obtained from the crystalline lens by passing a stream of carbon dioxide through a saline extract of this body.

Serum Globulins, constituents of blood plasma and blood serum.

Fibrinogen, which occurs in blood plasma and is converted into fibrin when the blood clots.

Myogen, a normal constituent of muscle.

Midway between these two groups may be placed the muscle protein, *myosin* (or *myosinogen*) which, though soluble in water, resembles the globulins in the ease with which it is precipitated by the addition of neutral salts.

Proteins allied to this class form an important constituent of plants, and are found in large quantities in many seeds used as articles of food. Prominent members of the vegetable globulins are the *edestins*, from hemp, cotton, and sunflower seeds, *zein* from maize, *legumin* from beans.

(5) **Gliadins,** contained in cereals, and soluble in dilute alcohol.

(6) **Glutelins,** also obtained from cereals, and soluble in weak alkalies or acids.

(7) **Derivatives of Proteins.** When proteins are heated with acids, or acted on at the body temperature by certain enzymes, *e.g.* pepsin, trypsin, or papain, they undergo hydrolysis. This action, when carried to its end, results in the production of the amino-acids which we have already dealt with.

These hydrolytic changes proceed by a series of stages, so that the intermediate products still present many of the protein reactions. These derived proteins are divided into three groups, *metaproteins*, *proteoses* and *peptones*. The formation of these intermediate products is especially marked with the proteolytic enzymes. Pepsin, together with hydrochloric acid, for example, only breaks down the protein molecule into smaller, but still quite large molecules, the proteoses and peptones. It does not liberate any amino-acids. Trypsin gives rise to proteoses, peptones and also amino-acids. In both cases, the peptide linkages are attacked, since equal numbers of NH_2 and COOH groups become free. The action of these enzymes on proteins, in fact, results in the production of a number of derivatives of progressively diminishing molecular weight and complexity by a hydrolytic process :—

$$\ldots \mathrm{R.CO.NH.R} \ldots \quad + \mathrm{H_2O} \longrightarrow \quad \ldots \mathrm{R.COOH} + \mathrm{NH_2.R} \ldots$$

The protein molecule is distinguished by the variety of the groups which enter into its formation, and this renders possible a great variety of intermediate products. Thus a protein molecule may consist of the amino-acids, A, B, C, D, E, F, G, H, etc., arranged in varying order. When hydrolysis occurs it may result in the immediate splitting off, say, of group A, while the residue breaks up into a series of proteoses whose composition may be represented as ABF, ABC, DFG, BDEF, etc. With further hydrolysis these groups are broken into still smaller ones, and the penultimate stages of the hydrolysis will be polypeptides.

No sharp dividing line can be drawn between the metaproteins, proteoses, peptones, and polypeptides. We have at present no satisfactory means of isolating the different members of these groups. Their classification is, therefore, a conventional one, depending on their solubilities and their precipitability by neutral salts. Both proteoses and peptones give reactions common to all proteins and, like these, are precipitated by such reagents as mercuric chloride, or phosphotungstic acid. They give the biuret reaction. Their solutions can be boiled without coagulation. Many of them may be thrown down from their solutions by alcohol.

(*a*) Metaproteins. These are amongst the early products of hydrolysis of the proteins and are little less complex than the proteins from which they are derived.

Acid Metaprotein is formed by the action of dilute acids on any proteins. If a weak alkali be added so as nearly to neutralise the solution of acid metaprotein, this latter is precipitated.

Alkaline Metaprotein is formed by the action of strong alkalies on any protein, or by adding alkali in excess to a solution of acid metaprotein. It is precipitated on neutralisation of its solution.

(*b*) PROTEOSES, *e.g. albumose* from albumin, *caseose* from casein. All of these are precipitated from their solutions on saturation with ammonium sulphate. In the presence of a neutral salt they give a precipitate on the addition of nitric acid. All, with the exception of heteroalbumose, are soluble in water, and all are soluble in weak salt solutions or dilute acids or alkalies. They are slightly diffusible through animal membranes.

(*c*) PEPTONES, *e.g.* fibrin peptone, gluten peptone. These are all soluble in water, diffuse fairly readily through animal membranes, and give reactions similar to those of the albumoses. From the latter they are distinguished by the fact that they are not precipitated on saturation of their solutions, either in acid or alkaline reaction, with ammonium sulphate. Many of them are soluble in alcohol.

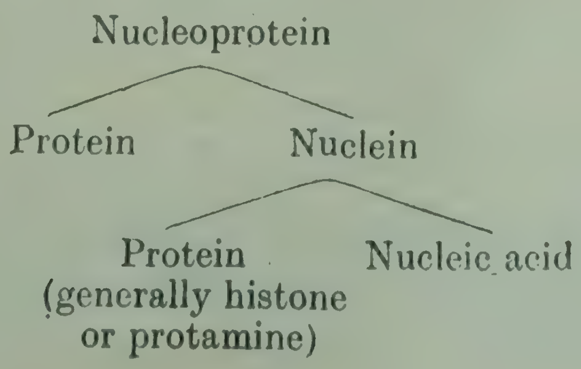
(*d*) THE POLYPEPTIDES resemble peptones, but are more readily diffusible.

(8) **Conjugated Proteins.** Various complex bodies which play an important part in building up cells, and in the vital processes of the body, make up this group of compounds. They resemble one another only in the fact that in each of them a protein molecule is combined with some other body, often spoken of as the *prosthetic* group.

(*a*) THE PHOSPHOPROTEINS. In this class may be included a number of substances of very diverse properties, which, however, resemble one another in containing phosphorus as an integral part of their molecule. The phosphoproteins have markedly acid characters. They are insoluble in water, easily soluble in alkalies, from which the original body is thrown down again on addition of acid. Their solutions in alkali are not coagulated by heating. To this class belong *caseinogen*, the chief protein of milk, and *vitellin*, the main protein in the yolk of egg. The vitellins are generally associated with a large amount of lecithin. The phosphoproteins differ from the nucleoproteins, which also contain phosphorus, in that they are readily decomposed by dilute caustic alkali with the liberation of phosphoric acid, and do not contain purine bases.

(*b*) CHROMOPROTEINS. Of this class, consisting of a colouring matter combined with a protein, the most important is *hæmoglobin*. This substance, which is the red colouring matter of the red corpuscles of the blood, is composed of the protein, globin, united with an iron-containing body, hæmatin $C_{32}H_{32}N_4O_4Fe$. Other very important chromoproteins are the *cytochromes*, which play an important part in the carrying out of tissue oxidations.

(*c*) THE NUCLEOPROTEINS. These are formed by the combination of a phosphorised organic acid, *nucleic acid*, with a basic protein which is usually one of the histones or protamines. If a nucleoprotein be subjected to gastric digestion, a large amount of the protein goes into solution as proteose or peptone, leaving an insoluble remainder. This precipitate still contains a protein group, and is spoken of as *nuclein*. From the latter nucleic acid can be split off by heating with stronger acids or by trypsin. The relation of the two portions of the nucleoprotein may be represented :—



The nucleoproteins are soluble in water, salt solutions and dilute alkalies. They have acid characters and are precipitated by acids. The nucleins, on the other hand, are insoluble in water and salt solutions, but are dissolved by dilute alkalies. Nucleoproteins form the chief constituents of cell nuclei, and the heads of spermatozoa therefore contain a large amount. They also occur in the cytoplasm. Nucleoproteins may be obtained from organs rich in cells, such as the thymus, and from yeast.

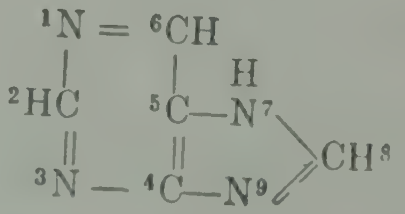
The nucleoproteins are of fundamental importance in all living things. The phenomena of heredity are closely connected with the properties of the nucleoprotein substance chromatin, and the hereditary elements called genes are probably molecules or aggregates of molecules of nucleoproteins. Also the viruses of disease, the simplest living things, are crystallisable nucleoproteins, which, like genes, can reproduce.

Nucleic Acid. By hydrolysis, the nucleic acid may be broken up into its proximate constituents. These differ slightly according to the source of the nucleic acid. Whatever the source, the disintegration products always contain phosphoric acid, purine bases, pyrimidine bases, and a pentose.

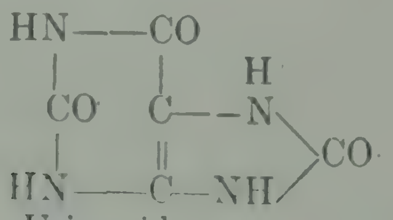
There appear to be two main types of nucleic acid, one present in nuclei, and therefore predominating in thymus nucleic acid, and one in cytoplasm, and especially abundant in yeast nucleic acid. These two differ slightly in their degradation products, yielding different pentoses and pyrimidines.

(1) *Phosphoric Acid.* The proportion of phosphorus in the nucleic acids is about 10 per cent. It represents, so to speak, the combining link for the groups contained in the nucleic acid molecule.

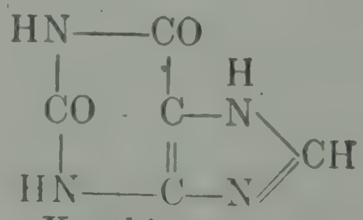
(2) *The Purine Bases.* Among the products of disintegration of nucleic acid we find constantly the bases *adenine*, $C_5H_5N_5$, and *guanine*, $(C_5H_5N_5O)$. These substances, with the products of their oxidation, *xanthine*, $C_5H_4N_4O_2$, *hypoxanthine*, $C_5H_4N_4O$, have long been known to be closely allied to uric acid, $C_5H_4N_4O_3$. According to Fischer they can be all regarded as derivatives of the body *purine* :



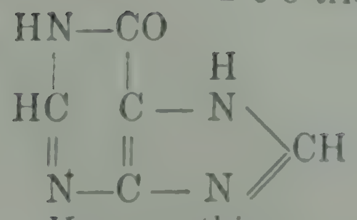
The relation of xanthine, hypoxanthine, guanine, and adenine to uric acid is shown by the following formulæ :



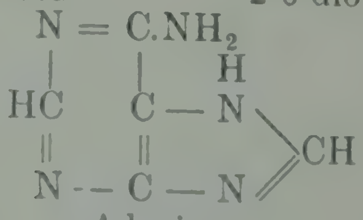
Uric acid
2-6-8-trioxypurine



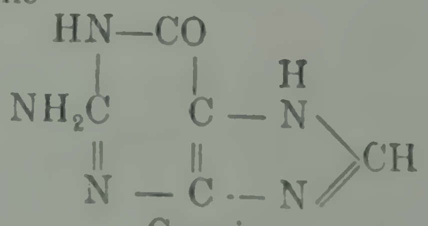
Xanthine
2-6-dioxypurine



Hypoxanthine
6-oxypurine



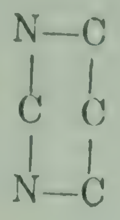
Adenine
6-amino-purine



Guanine
2-amino 6-oxypurine

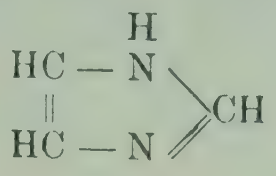
Closely allied to this group of bodies are the active constituents of tea, coffee, and cocoa, namely *caffeine*, which is *trimethyl dioxypurine*, and *theobromine*, which is *dimethyl dioxypurine*.

From the structural formulæ given it will be seen that the purine ring contains two nuclei. The nucleus :



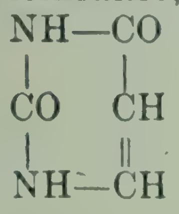
is spoken of as the *pyrimidine nucleus.*

The other is the *iminazol* ring which we have met with already in histidine :

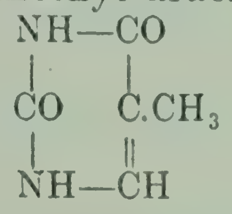


(3) *Pyrimidine Bases.* In addition to the pyrimidine ring contained in the purines, nucleic acids yield a separate series of bases derived from the pyrimidine ring. These are uracil, thymine, and cytosine.

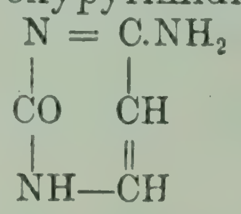
Uracil, 2-6-dioxy-pyrimidine,



Thymine, 5-methyl uracil,



Cytosine, 6-amino-2-oxypyrimidine,



Nucleic acids from nuclei yield cytosine and thymine ; the acids from cytoplasm give cytosine and uracil.

(4) *Pentose.* In the nucleic acids of nuclei this is a desoxypentose ($C_5H_{10}O_4$), viz. *β-d*-2-desoxyribofuranose, but in nucleic acids of cytoplasm and in the mononucleotides of animal tissues, *e.g.* adenylic acid, the carbohydrate group is the ribose, *β-d*-ribofuranose of empirical formula $C_5H_{10}O_5$. In general, the two nucleic acids comprise respectively the following units :

Nuclear.	*Cytoplasmic.*
Phosphoric acid.	Phosphoric acid.
Desoxyribose.	Ribofuranose.
Adenine and guanine.	Adenine and guanine.
Cytosine and thymine.	Cytosine and uracil.

The researches of Levene and of others have thrown light on the manner of combination of these groups in the molecule of nucleic acid. It appears that each purine or pyrimidine group can combine with a pentose to form what is called a "*nucleoside,*" *e.g.* :—

$$\underset{\text{(Adenine)}}{C_5H_4N_5}—\underset{\text{(Ribose)}}{C_5H_9O_4} \text{ is the nucleoside } \textit{adenosine.}$$

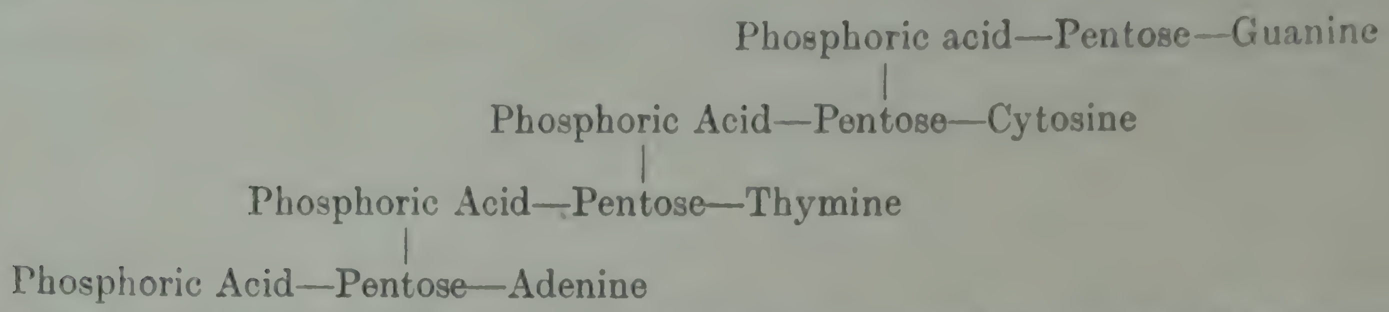
Each nucleoside can be then attached to a molecule of phosphoric acid, this complex being named a "*nucleotide*" and representing a nucleic acid in its simplest form, *e.g.* :—

$$\underset{\text{(Adenine)}}{C_5H_4N_5}—\underset{\text{(Ribose)}}{C_5H_8O_3}—\underset{\text{(Phosphoric acid)}}{PO_3H_2} \text{ is the (mono) nucleotide } \textit{adenylic acid,}$$

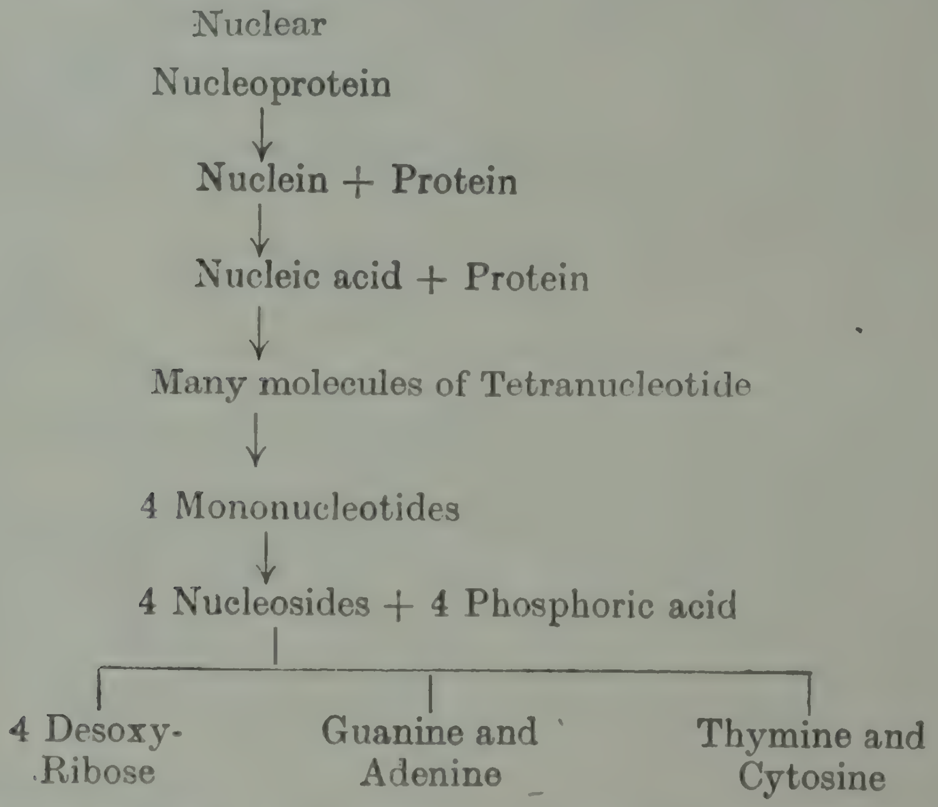
or *adenosine monophosphoric acid.*

By addition of one and two further molecules of phosphoric acid, we get the *adenosine diphosphoric* and *triphosphoric* acids, which play most important functions as carriers of phosphoric acid radicles in many biochemical processes. Other important nucleotides are *nicotinic amide mononucleotide, flavine mononucleotide* and *adenine-flavine dinucleotide,* all coenzyme components.

In general, nucleic acids are very large molecules, the molecular weight of th smallest probably being about 15,000; these are aggregated into elongate particles of weights ranging from 200,000 to 1,000,000. In structure, th smaller units appear to be chains of relatively small molecules called *tetr nucleotides*, in which four mononucleotides are linked together, possibly throug the phosphoric acid molecule of one and the carbohydrate of the next. Thu thymus nucleic acid would be built up from units, each of them a tetranucleotid which would have the structure


Phosphoric acid—Pentose—Guanine
Phosphoric Acid—Pentose—Cytosine
Phosphoric Acid—Pentose—Thymine
Phosphoric Acid—Pentose—Adenine

The degradation of nucleoprotein may be rendered clearer by the following scheme:


Nuclear
Nucleoprotein
Nuclein + Protein
Nucleic acid + Protein
Many molecules of Tetranucleotide
4 Mononucleotides
4 Nucleosides + 4 Phosphoric acid
4 Desoxy-Ribose
Guanine and Adenine
Thymine and Cytosine

(*d*) THE GLYCOPROTEINS. In the glycoproteins the prosthetic group is represented by a carbohydrate radical, generally containing nitrogen, such as glucosamine or galactosamine. Serum albumin and globulin both contain a carbohydrate consisting of glucosamine and mannose.* The glycoproteins proper are split into their two constituents, protein and carbohydrate, on boiling with dilute mineral acids. The carbohydrate prosthetic group is a tetrasaccharide, derived *e.g.* from two acetylhexosamine molecules joined together by two glucuronic acid molecules. The hexosamine called glucosamine gives a tetrasaccharide called *mucoitin*, while galactosamine yields the tetrasaccharide called *chondroitin*—they occur respectively as mucoitin and chondroitin sulphuric acids. Heparin is a derivative of some such compound as chondroitin sulphuric acid.

The glycoproteins are divided into the two main groups, viz.: mucins and mucoids.

The *mucins* are present in the slimy secretion of the mucous membranes and which is secreted either by goblet cells or by mucous glands. They may be precipitated from their solutions by the addition of acids, and redissolve on addition of alkalies. They are not coagulable by heat. They give the various typical protein tests, such as the xanthoproteic, Millon's, the biuret reaction and so on. Prolonged boiling with acids splits the molecule, with the production of acid metaprotein and mucoitin sulphuric acid.

---

* RIMINGTON. *Biochem. J.*, 1931, **25**, 1062.

The *mucoids* include a number of substances which may be extracted from various tissues by the action of weak alkalies, *e.g.* from tendons, bone and cartilage. Another example of this class of mucoids is *ovomucoid*, which is a constituent of egg-white. A similar body has been prepared from blood serum. The best studied example of this group is the *chondromucoid* which, with collagen, forms the ground substance of cartilage. Chondromucoid on hydrolysis yields a protein, and both mucoitin and chondroitin sulphuric acids.

(9) **The Albuminoids or Sclero-proteins.** Under this heading are grouped a number of diverse substances which play an important part in building up the framework of the body. They play the part in the animal economy which in the vegetable kingdom is filled by the celluloses, lignin, the pentosans, etc. Their value as skeletal tissue seems to be determined by their insoluble character, but on this account it is practically impossible to purify them ; we can simply take the residue of a tissue which is left after extraction of the soluble constituents. When broken down by the action of strong acids, they yield disintegration products which resemble those derived from other proteins.

Collagen forms the main constituent of white fibrous tissue and the ground substance of bone and cartilage. Under the action of acids or when subjected to prolonged boiling with water it is converted into *gelatin*, which forms a colloidal solution setting to a jelly when cold. When subjected to acid hydrolysis it gives a series of amino-acids from which tyrosine and tryptophane are wanting. On the other hand, there is a preponderance of glycine and phenyl alanine. Gelatin is hydrolysed by gastric juice or by pancreatic juice, whereas collagen is unaffected by the latter.

The *Keratins* are produced by the modification of epithelial cells, and form the horny layer of the skin as well as the main substance of hairs, wool, nails, hoofs, horns and feathers. They are distinguished by their insolubility in water, dilute acids or alkalies, and if eaten pass through the alimentary canal unchanged. Although differing in their elementary composition, they are all distinguished by the very large amount of sulphur present in their molecule. The greater part of this sulphur is in the form of cystine, of which keratin contains as much as 10 per cent.

Neurokeratin, which forms the basis of the neuroglia of the central nervous system, is grouped with the keratins. It is extracted from nervous tissues by boiling these with alcohol and ether, and then submitting the tissue to tryptic digestion which leaves the neurokeratin unaffected.

Elastin forms the elastic fibres of the connective tissues. Elastin is insoluble in water, alcohol or ether, or in dilute acids and alkalies. It is slowly dissolved on *prolonged* treatment with gastric juice, but is practically unaffected in the alimentary canal.

Other members of this group are *fibroin*, which forms the main substance of silk; *spongin*, the horny framework of sponges ; *conchiolin*, the ground substance of shells : and perhaps the amyloid substance which we have already mentioned in connection with the mucoids. All these sclero-proteins present considerable differences in their qualitative and quantitative content of amino-acids.

Specificity of Proteins. Although certain proteins may appear to be identical so far as ordinary chemical and physical properties, and even so far as chemical analysis is concerned, they are not necessarily identical in their biological properties. For example, casein from cow's milk and casein from goat's milk, or egg albumin from hen or duck, have certain specific properties by which they are distinguished from one another. This can be shown by *anaphylactic reactions* (*v.* Chapter XXXVII.). Thus, if a guinea-pig is sensitised by injections of hen-egg albumin, administration of egg albumin from another species is not followed by any effects, whereas an extremely small injection of the hen albumin promptly causes death from anaphylactic shock. There is thus a species specificity with proteins which is not seen with other biochemical constituents.

Dudley and Woodman * showed that this is in all probability due to the fact that in the different species there are slight differences in the *arrangement* of the constituent amino-

---

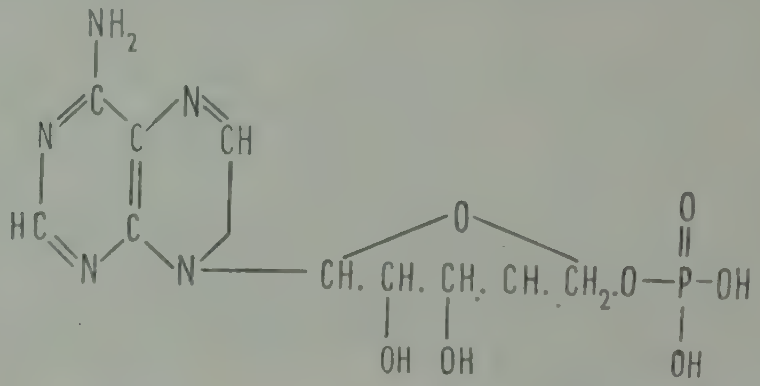
* Dudley and Woodman. *Bioch. Journ.*, 1915, 9, 97.

OOD TECHNOLOGICAL RESEA

acids in the protein molecule. Thus if in one species we have the amino-acids A, B, C, D, . . . arranged in that order, we should have a stereochemically different molecule if t arrangement were B, D, A, C, E . . . . These species differences are of great importance i connection with the subject of immunity; they also explain why, in the construction protein in the body, the protein taken as food must first be thoroughly broken down int its constituent amino-acids.

SOME IMPORTANT PHOSPHORIC ACID DERIVATIVES

It will be useful at this stage to mention some further compounds containin phosphoric acid in the molecule, which play a most fundamental *rôle* in almos all the chemical reactions concerned with energy liberation. The outstandin substance in this respect is *adenosine triphosphate*. The adenosine phosphate are all derived from the nucleoside *adenosine* (p. 49), a compound of th amino-purine adenine with the pentose *d*-ribose; it can combine with one two or three molecules of phosphoric acid to give adenosine mono-, di- o triphosphate, as shown below :—



(*Adenine*) (d-*ribose*) (*Phosphoric acid*)

Adenosine monophosphate

or, for brevity $\textcircled{Ad}$—Ph, where $\textcircled{Ad}$ stands for adenosine and Ph for $-PO_3H_2$

Similarly the diphosphate and triphosphate may be written briefly :—

$\textcircled{Ad}$— Ph $\sim$ Ph *Adenosine diphosphate* (A.D.P.)

$\textcircled{Ad}$— Ph $\sim$ Ph $\sim$ Ph *Adenosine triphosphate* (A.T.P.)

Now the importance of these compounds lies in the fact that the second and third phosphoric groups are bound to the monophosphate phosphoric group by what are called *energy-rich* bonds, indicated above by the wavy bond $\sim$. This means that when these groups are split off a large amount o energy is set free. If the bond is broken by hydrolysis the energy appears as heat and is equal to 12 Cal. per g. molecule :—

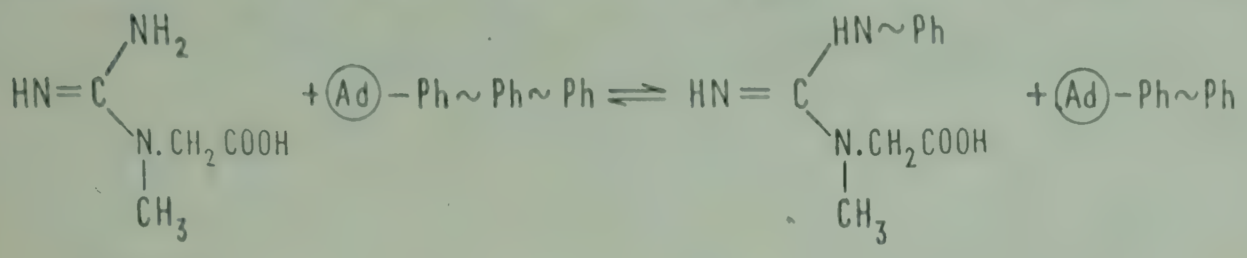
$$\textcircled{Ad}-Ph \sim Ph \sim Ph + H_2O \longrightarrow \textcircled{Ad}-Ph \sim Ph + H_3PO_4 + 12\ \text{Cal.}$$

But what is even more important, 10 Cal. of the energy set free by the liberation of the phosphate group is free energy, *i.e.* it can be directly utilised to do work instead of being converted into heat, and it is probable that all the energy manifestations of living things are derived from the resolution of such energy-rich bonds. One of the uses to which living cells can put such energy is to synthesise new compounds of high energy potential, in which, so to speak, energy is stored up ready for use. It is especially in these syntheses that the energy-rich phosphate groups are concerned, because in such reactions the phosphate group can be added on to other compounds, together with the energy-rich bond, without loss of energy.

Another important phosphoric acid derivative is *creatine phosphate*.* It also contains an energy-rich phosphate group, and its formation is an example of the transference of this group, without energy loss, from A.T.P.

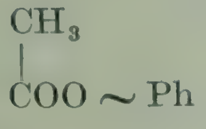
---

* Eggleton. *Physiol. Rev.* 1929, **9**, 432.



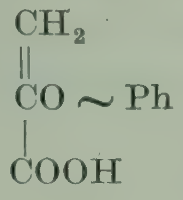
(*Creatine*) (*Creatine phosphate*)

Other compounds with energy-rich phosphate bonds are phosphoacetic acid, phospho-*enol*-pyruvic acid, and 1 : 3 diphosphoglyceric acid.



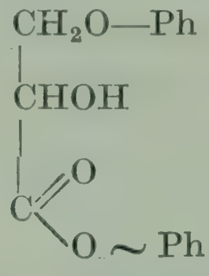
$$\begin{array}{l} CH_3 \\ | \\ COO \sim Ph \end{array}$$

Phosphoacetic acid



$$\begin{array}{l} CH_2 \\ \| \\ CO \sim Ph \\ | \\ COOH \end{array}$$

Phospho-enol-*pyruvic acid*



1 : 3 *Diphosphoglyceric acid*

We shall see later that these various compounds play an important *rôle* in the development of the energy of muscular contraction, and in processes of intermediary metabolism.

For Reference

Astbury (1933). Fundamentals of Fibre Structure. (Oxford : Univ. Press.)

Chibnall (1942). Amino-acid Analysis and the Structure of Proteins. *Proc. R. S.*, **131B**, 136.

Cohn (1938). Proteins as Chemical Substances and as Biological Components. (*Harvey Lects.*, 1938–9 ; p. 124.)

Danielli and Brown, *Ed.* (1946). Nucleic Acid. *Symposia Soc. Exp. Biol.* (Cambridge Univ. Press.)

Kossel (1928). The Protamines and Histones. (Longmans, Green & Co.)

Levene (1925). Hexosamines and Mucoproteins. (Longmans, Green & Co.)

Levene and Bass (1931). Nucleic Acids. *Amer. Chem. Soc. Monograph*, **56**.

Lloyd, Dorothy Jordan and Shore, Agnes (1938). Chemistry of the Proteins. (London, Churchill.)

Mitchell and Hamilton (1929). Biochemistry of Amino-Acids. *Amer. Chem. Soc. Monograph, No.* **48**.

Osborne (1924). The Vegetable Proteins. (Longmans, Green & Co.)

Royal Society (1939). A Discussion on the Protein Molecule. *Proc. R. S.*, **127B**.

Taylor (1937). Chemistry of Proteins and Related Substances. *Ann. Rep. Prog. Chem.*, **34**, 302.

Vickery and Osborne (1928). A Review of Hypotheses of the Structure of Proteins. *Physiol. Rev.*, **8**, 393.

C. BIOPHYSICAL PRINCIPLES *

CHAPTER V

THE ENERGY OF MOLECULES IN SOLUTION

WATER forms from 70 to 95 per cent. of the mass of protoplasm. No substance introduced into the protoplasm has any chemical influence on it unless it be soluble, and the first stage in the preparation of foodstuffs for assimilation always consists in a process of solution. It follows that all the energy with which we have to deal is the energy of *molecules in, or in equilibrium with, watery solution.*

DIFFUSION. When a substance (solute) is dissolved in a liquid (solvent), we may regard the molecules of both solute and solvent as being in continual movement within the liquid, very much as the molecules of a gas are in continual motion. In consequence of this state of movement solutes travel by diffusion from places where they are more concentrated to parts where they are more dilute, until finally a uniform concentration obtains throughout the solution. Thus, if a vessel be half filled with a 10 per cent. solution of cane sugar, and be then filled up carefully with water, so as to form a layer on the heavier sugar solution, the sugar gradually diffuses up into the water, and after some weeks there will be the same concentration of sugar throughout all parts of the fluid. Carried out in this way, the process of equalisation is a slow one, but when aided by the production of currents, or when only thin films of liquids are concerned, spread by diffusion may be effected in quite short periods of time.

We also find differences between the rates of diffusion of dissolved substances —determined by the weight and size of the individual molecules, although the relation between molecular weight, or size, and rate of diffusion is not so simple as in gases.

The diffusibility of a substance is given by its *diffusion coefficient.* If we take a cylindrical mass of solution which is one centimetre long and has a sectional area of one square centimetre (Fig. 14), and maintain a constant difference of concentration between A and B = 1, the diffusion coefficient is the amount of substance which diffuses in a unit of time from A to B. Thus the statement that the diffusion coefficient of urea is 0·810 at 7·5° C. denotes that, if A be continually filled with a 1 per cent. solution of urea, while in B a constant current of distilled water is kept up so as to maintain the concentration at zero, in the course of a day, at 7·5° C., 0·810 gramme of urea will pass from A to B through the cylinder of one centimetre in length and one square centimetre in cross section. The determination of these diffusion coefficients is rendered easier by the fact, first ascertained by Graham, that diffusion of salts occurs as rapidly through a solid jelly of gelatine or agar-agar as through water. It is therefore possible to make the cylinder in the diagram solid by the admixture of one of these two substances, and to maintain a constant concentration on the two sides of it by the circulation of fluid without affecting the rate of diffusion through the cylinder by setting up accidental currents.

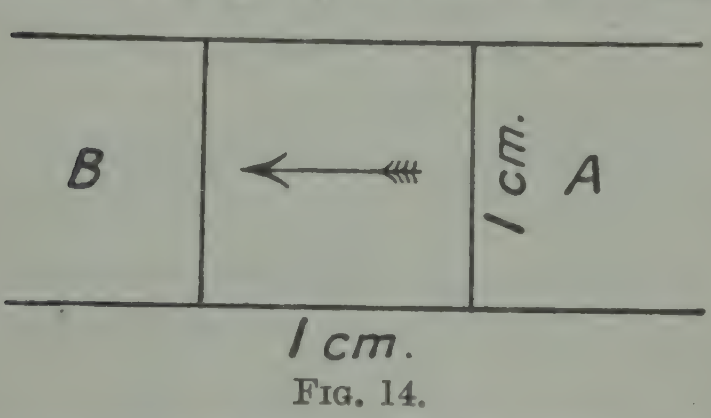


FIG. 14.

Filtration, Ultra-filtration and Dialysis. Filter paper permits the passage both of water and of all dissolved substances. Different kinds of filter paper allow passage with varying degrees of rapidity, depending on the fineness of the holes through which the fluid can pass. The same difference may be observed in porous structures, such as the Berkefeld filter, which allow all dissolved substances to pass but keep back most bacteria. On the other hand the very smallest micro-organisms, the so-called filter-passers, will traverse the pores of a Berkefeld filter. Most gels, such as gelatin and collodion are of porous structure, and collodion membranes may be obtained of varying degrees of permeability according to their mode of preparation * and the concentration of the collodion in the membrane. Such membranes are used for the process known as *ultra filtration*. The membranes of loosest texture will allow proteins such as albumin to pass, in addition to crystalloids. On the other hand the denser collodion membranes, or cellophane, will keep back all colloid constituents and allow only water and crystalloids to pass. If, for instance, serum be placed on one side of such a membrane, under pressure a filtrate will be obtained containing the water, salts, urea and other crystalloid constituents of the serum (Fig. 15).

---

* As with the chapters on Structural and Biochemical Principles, the present chapter can hope to be no more than an outline indicative of the main aspects of the important branch of biology called Biophysics, for further details of which special treatises must be consulted.

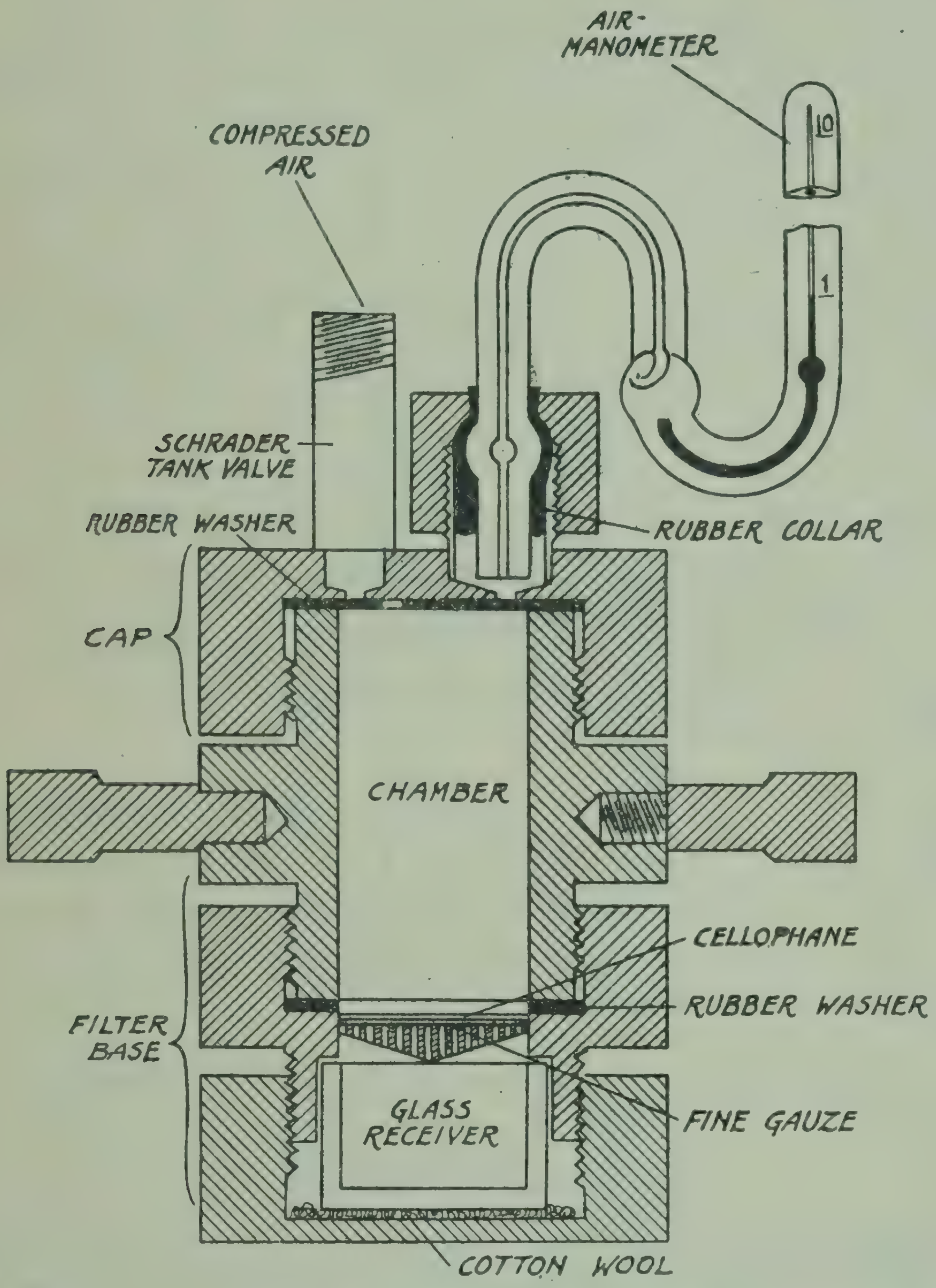


FIG. 15. Apparatus for ultra-filtration. The filtering membrane is cellophane. The upper chamber contains the colloid solution, from which a colloid-free filtrate is forced through into the lower chamber by air pressure. (AITKEN and KAY, *J. Physiol.*, 1927, **62**, Proc. xxix.)

Similar membranes are also made use of when it is desired to free a colloidal solution from salts, in the process known as *dialysis*. If a fluid containing both colloids and crystalloids in solution, *e.g.* blood serum, be enclosed in a tube of vegetable parchment or collodion which is hung up in a large bulk of distilled water (Fig. 16), all the salts diffuse out, and if the water be frequently changed, we obtain finally a fluid within the "dialyser" free from salts and other crystalloid substances, but containing the whole of the colloids originally present. In the process of dialysis, however, not only does solute pass out into the water, but water also passes inwards into the solution, so that this tends to become diluted. These phenomena are easily explained if we remember that molecules of solvent and solute are in constant movement, and if we regard the solution as being weaker in solvent than the pure solvent is. Thus, to consider a strong cane sugar solution and water or a weak cane sugar solution, the former consists of a high concentration of sugar molecules and a lowered concentration of water molecules, while the latter is a high concentration of water molecules and a low (or non-existent) concentration of sugar molecules. Consider that we have a vessel with two such "solutions," A strong and B weak (or B = water), at the same level, and separated by a membrane *pm*, permeable to *both* sorts of molecule, as in Fig. 17 (I). Then solvent and solute can both

---

* PIERCE. *J. Biol. Chem.*, 1927, **75**, 795; DESAI, *Bioch. J.*, 1930, **24**, 1897.

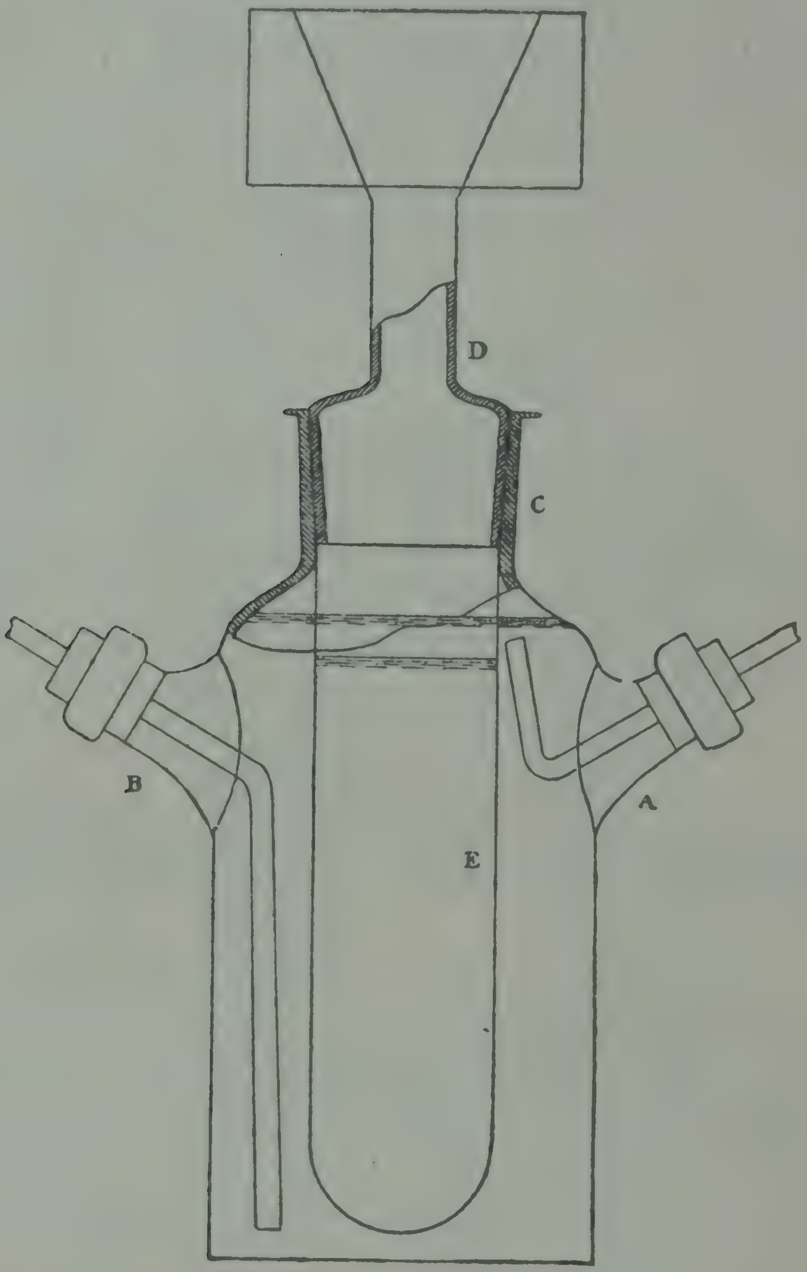


Fig. 16. Dialysing Apparatus. The collodion tube E is fastened to the glass stopper D which fits airtight into the neck of the bottle. Sterile water is introduced into the bottle from a reservoir through B, and the colloid for dialysis poured into E. A partial vacuum is then made in the bottle by exhaustion at A; this opposes the osmotic pressure of the solution in E and so prevents its dilution, or if great enough may even concentrate it by ultra-filtration. (Jessne Hansen.)

move in either direction, but because solute is more concentrated in A and solvent more concentrated in B, there will be a greater passage of solute into B, and of solvent into A, than in the reverse direction. Ultimately, when the two solutions have become uniform in composition, A by losing solute and gaining water, B by gaining solute and losing water, there would at first be a raised level of solution in A as compared with B, Fig. 17 (II). But this difference of hydrostatic pressure, since the membrane is freely permeable to both solvent and solute, causes fluid of the uniform composition to flow from A to B, until it reaches the same level, and true equilibrium is now reached (Fig. 17 (III)).

If a certain pressure is put on the contents of the dialyser, or (as in Fig. 16) suction on the outside, fluid also passes out by ultra-filtration, but there is a certain limiting

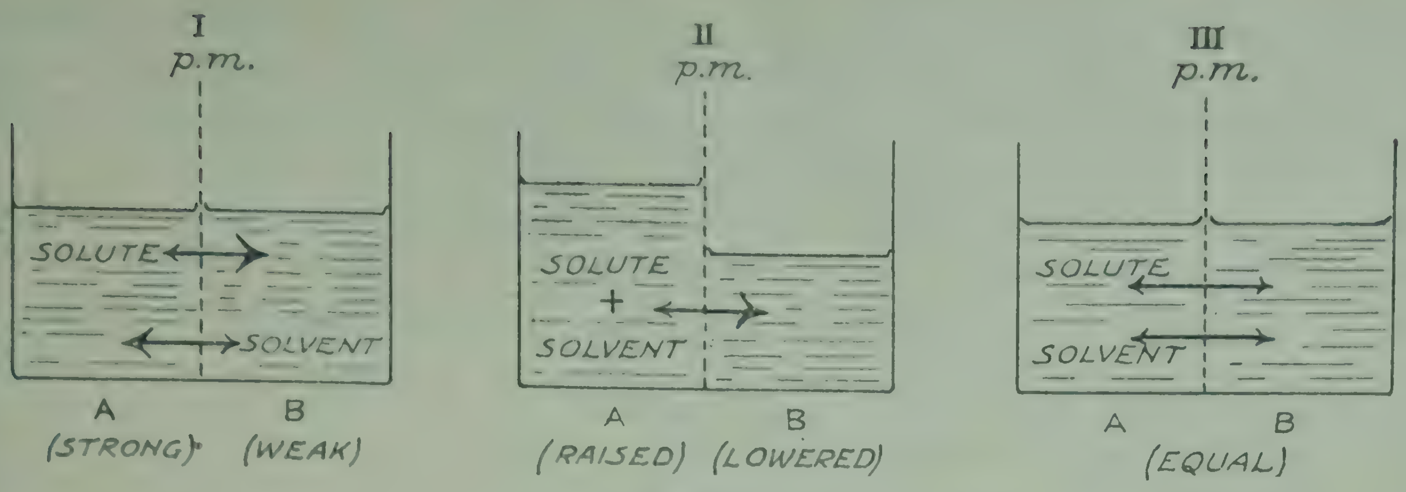


FIG. 17.

pressure below which no filtration will take place. This limiting pressure is equivalent to the osmotic pressure of the colloids in solution, and varies according to the state of aggregation of the colloid, *i.e.* to the number of molecules, simple or compound, of the colloid which is present. In blood serum it is equal to about 30 mm. Hg. If the difference of pressure on the two sides of the membrane is less than this amount, water and salts will pass from the outside to the inside of the membrane and the serum will be diluted.

Semi-permeable Membranes. Membranes may be of different degrees of permeability, according to the sizes of their structural pores, so that some solutes may be let through, but not others of larger molecular size. A special case arises when only the solvent can pass through, while all solutes are held back. This is called a *semi-permeable membrane.* If we consider the case of a solution separated from its solvent by such a membrane, it is clear that if we start with pure solvent on one side, equilibrium can never be established in the sense that the solutions will become of the same composition on both sides of the membrane, because ultimately all the solvent could be absorbed into the solution, since no solute would pass out.

OSMOTIC PRESSURE

This attraction of solvent to a stronger from a weaker solution is known as *osmotic pressure*, or, more correctly, as *osmotic water attraction.* Consider the case of a strong and a weak solution (or pure solvent) separated by a semi-permeable membrane, *s.p.m.* in the vessel A, B (Fig. 18), each limb of which is provided with a watertight piston, bearing a scale-pan *a*, *b*. Left to itself, solvent will pass from B to A, owing to the osmotic pressure, or osmotic water attraction, of the solution in A. If a weight W, be placed on the pan *b*, the rate at which water passes into A will be accelerated, because solvent will be driven through the *s.p.m* by the pressure put upon it. If, instead, the weight be placed on the pan *a*, as in the figure, the rate at which solvent enters A will be reduced, and if the weight placed on *a* be heavy enough, solvent can be driven out of A into B, against the osmotic water attraction of the solution in A. If the weight on *a* be just sufficient to prevent water either entering A or leaving it, a state of equilibrium is reached, the weight just counteracting the osmotic pressure of the solution, and so giving a measure of its magnitude.

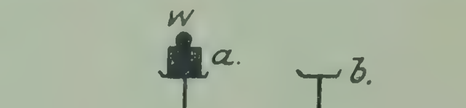


FIG. 18.

Osmotic pressures can only be set up permanently by solutes to which the membrane is impermeable, while permeable to the solvent.

Direct experiment has shown that the osmotic pressure exerted by a solute in a given volume of solvent is the same as the pressure the same weight of solute would exert when in a state of gas occupying the same volume as the solvent, *i.e.* one gramme molecule of solute in 22·4 l. of solvent would have an osmotic pressure of one atmosphere (or 760 mm. Hg). We see that the osmotic pressure of a solution, the measure of its attraction for water, represents a certain amount of potential energy, which can be utilised in an osmotic machine, such as that represented in the diagram, for the performance of work.

The amount of work done in this way can be calculated in exactly the same way as when we are dealing with a gas. To compress a gramme molecule of any perfect gas from volume V_1 to Volume V_2 requires the expenditure of work represented by $W = RT \log e \frac{V_1}{V_2}$. If we take the pressures, and compress from pressure p_1 to pressure p_2, the expression will be $W = RT \log e \frac{p_2}{p_1}$. To convert from natural to ordinary logarithms, we must multiply this expression by 2·303. The numerical value of R depends on the units required, being 1·991 for gramme calories and 0·848 for kilogramme metres. T at 37° C. is 310°. The equation then becomes :

$$\left.\begin{array}{r} W = 1421 \text{ gramme calories} \\ \text{or } 605{\cdot}5 \text{ kilogramme metres} \end{array}\right\} \times \log\frac{p_2}{p_1}.$$

To compress a fluid containing one gramme molecule of solute from osmotic pressure π_1 until it is concentrated to an osmotic pressure π_2, would require $605{\cdot}5 \times \log \frac{\pi_2}{\pi_1}$ kilogramme metres.

It must be remembered that this expression gives the work done in concentrating a solution, whatever method of concentration be applied, (*e.g.* evaporation) at a constant temperature.

THE MEASUREMENT OF OSMOTIC PRESSURE. By a method differing but little from the one just sketched out, it is possible directly to measure the osmotic pressure of aqueous solutions. A semi-permeable membrane is provided when copper ferrocyanide is deposited within the interstices of a clay cell. The cell is then filled with the required solution, connected with a manometer and immersed in distilled water, which passes into the cell until the pressure within the latter is equal to the osmotic pressure of the dissolved substances. By this means Pfeffer obtained the results given in the table below with a 1 per cent. solution of cane sugar at different temperatures.

By Avogadro's hypothesis, equal volumes of gases at the same pressure contain equal numbers of molecules. The figures show that the same laws apply to the osmotic pressure of substances in solution as to the pressure of gases in their free state. It is therefore possible to reckon the osmotic pressure which would be exerted by 1 per cent. sugar in solution at a given temperature.

Temp. °C.	Pressure in atmospheres.	
	Observed.	Calculated.
	Atm.	Atm.
6·8	0·664	0·665
13·7	0·691	0·681
22·0	0·721	0·701
32·0	0·716	0·725
36·0	0·746	0·735

This calculation is carried out as follows : A gramme molecule of any gas at 0° C. and 760 mm. Hg has a volume of 22·4 litres, therefore 342 grammes of cane sugar (one mole), if it could be converted into a gas at 0° C. and 760 mm. Hg, would have a volume of 22·4 litres. In Pfeffer's experiment one gramme of sugar was dissolved in 100 grammes of water making a total volume at 0° C. of 100·6 c.c. The osmotic pressure of the sugar molecules in this solution will therefore amount to $\frac{22400}{342 \times 100{\cdot}6} = 0{\cdot}651$ atmosphere. At a temperature of 6·8° the pressure would be 0·665 atmosphere, as against the observed 0·664 atmosphere.

Pfeffer's method is difficult to carry out and is not applicable to all dissolved substances. Other indirect methods have therefore been applied to the comparison of the osmotic pressures of different solutions.

Determination by Plasmolysis. Solutions which have the same osmotic pressure are spoken of as isosmotic. The method of plasmolysis, which we owe to the botanist De Vries, consists essentially in the comparison of the osmotic pressure of solutions with that of the cell sap of certain plant cells, and depends on the principles already explained viz.: that the superficial layers of cytoplasm behave to many solutes as semi-permeable membranes. If we determine the concentration of a solution at which the *plasmolysis* just occurs, and another smaller concentration at which plasmolysis is absent, we know that the concentration of the cell sap lies between those of the two salt solutions. Thus, if plasmolysis occurs in a solution containing 0·6 per cent. sodium chloride and is absent in a solution containing 0·59 per cent. of the same salt, the concentration of the cell sap must be about equivalent to a 0·595 per cent. NaCl solution. Solutions of different salts, in which plasmolysis just occurs, must also be isotonic with one another. Thus a 1·01 per cent. solution of KNO_3 is found to be isotonic with a 0·58 per cent. NaCl solution.

Determination by Hamburger's Blood Corpuscle Method. The external layer of red blood corpuscles is also impermeable to a number of dissolved substances. If red corpuscles be placed in a solution of such a substance of smaller osmotic pressure than their contents, they will swell up and, since they have no supporting cell wall, the increase in size may go on until the corpuscles burst, and their contained red colouring matter, hæmoglobin, passes into solution in the surrounding fluid. If the corpuscles be then allowed to settle or be centrifuged, the fact that hæmolysis has occurred is shown by the red colour of the clear supernatant fluid. In order to determine the osmotic pressure of a solution, say of sugar or of sodium chloride, these are added in various dilutions to blood corpuscles until we get solutions in which hæmolysis just occurs. These are called *isotonic* solutions. As an example of this method may be adduced the following results :

	Lowest concentration in the solution in which the blood corpuscles do not hæmolyse	Highest concentration of the solution in which the blood corpuscles begin to hæmolyse	Mean concentration
	Per Cent.	Per Cent.	Per Cent.
Potassium nitrate . .	1·04	0·96	1·00
Sodium chloride . .	0·60	0·56	0·585
Cane sugar . . .	6·29	5·63	5·96
Potassium iodide . .	1·71	1·57	1·64
Sodium iodide . . .	1·54	1·47	1·505
Potassium bromide . .	1·22	1·13	1·17

A more accurate approximation can be got by finding the concentration of the given substance at which the red corpuscles remain unaltered in dimensions or in volume.

Isotonic is not necessarily identical with isosmotic. Tonicity has reference to the action of the solution with regard to some cell membrane, and therefore implies the impermeability of the membrane to the substance or substances in solution. Thus equimolecular solutions of urea and of glucose would be isosmotic, yet, tested by the blood corpuscle method, the solution of urea would be found to be isotonic with distilled water, since the outer layer of a blood corpuscle is freely permeable to urea, but not to sugar. Urea is therefore unable to exert any osmotic pressure in relation to blood corpuscles.

Indirect Methods of Measuring Osmotic Pressure. Equimolecular solutions (of non-electrolytes) have the same osmotic pressures. Since the osmotic pressure of such a solution is dependent on the number of molecules of solute in unit volume, any method which will tell us this will also enable us to determine the osmotic pressure. Other properties of solutions which are functions of the number of molecules present, are vapour tension, boiling point, freezing point. The presence of a substance in solution in water diminishes its vapour tension at any given temperature, raises its boiling point, and depresses its freezing point; and the extent of the deviation is proportional to the number of dissolved molecules present. The determination of the rise of boiling point is of very little value in physiology.

The classical method is the measurement of the *depression of freezing point* (generally represented by $\triangle$).

The determination is carried out with a thermometer reading to $\frac{1}{100}$° C. (Fig. 19). The freezing point of a decimolar solution of a non-electrolyte is $-0{\cdot}186°$ C., and that of a molar solution (*i.e.* containing one gramme molecule per litre, in the case of glucose 180 grammes), would have a $\triangle$ of 1·86° C. Since this has, as already explained, an osmotic pressure of 22·4 atmospheres, it is evident that the depression of freezing point can be converted directly into osmotic pressure by multiplying the depression by $\frac{22{\cdot}4}{1{\cdot}86} = 12{\cdot}04$. Thus a 1 per cent. solution of sodium chloride with $\triangle = 0{\cdot}61$ will have an osmotic pressure of $0{\cdot}61 \times 12{\cdot}04 = 7{\cdot}34$ atmospheres, while a 1 per cent. cane sugar solution with a $\triangle = {\cdot}0546$

will have an osmotic pressure of 0·675 atmosphere. This method has the advantage that the fluids are in most cases not altered by freezing. For colloid solutions, since the osmotic pressures are small, the values of $\triangle$ are very small.

A simple and accurate method was introduced by A. V. Hill *; it depends on the determination of the *depression of vapour pressure* by the dissolved substances, by observation of the differences in temperature produced by evaporation. Since it involves the use of thermopiles it is extremely delicate. The method makes use of the fact that if two aqueous solutions of different osmotic (*i.e.* vapour) pressures are placed in a moist atmosphere, water will be evaporated faster (or condensed more slowly) at the surface of the solution having the lower than from the one with the higher osmotic pressure, and hence the solution with the lower osmotic pressure will become cooler than the other. If the two solutions are in contact with the two faces of a thermopile connected with a galvanometer, the difference in temperature produces a deflection proportional to the differences in osmotic pressure. If the osmotic pressure of one of the two solutions is known, that of the other can be deduced. In practice a piece of filter paper is wetted with each of the two solutions, *e.g.* blood on one and 0·92 per cent. NaCl on the other, and these are applied one to each face of the thermopile. This is then placed in a moist chamber kept damp by the known solution and at constant temperature, and connected with the galvanometer. Readings are taken after half an hour. Example: Blood and 0·92 per cent. NaCl. Galvanometer reading = 19·8 mm. Calibration showed that 1 mm. deflection = ·001120 grms. NaCl per cent. Hence difference of V.P. on two sides is = 19·8 × ·001120 = ·0222 grms. per cent., *i.e.* the V.P. and O.P. of blood is equal to that of 0·92 + ·0222 = 0·9422 grms. NaCl per cent.

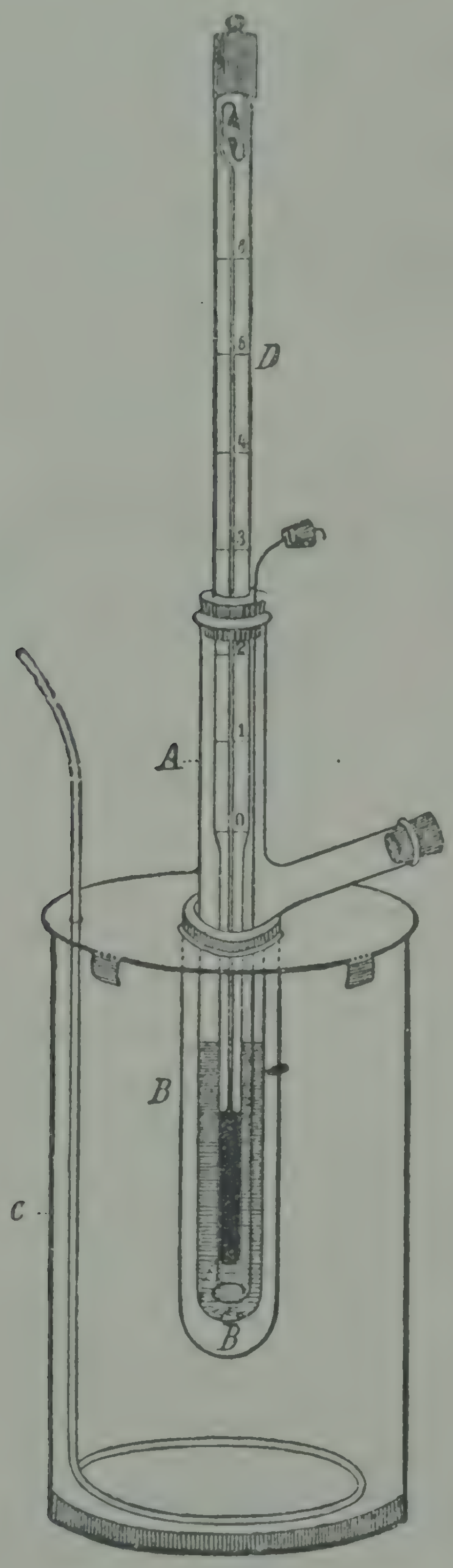


FIG. 19. Beckmann's Apparatus for Determination of Freezing Point.

---

* HILL. *Proc. Roy. Soc. A.*, **1930, 127, 9.**

Barger suggested a simple method * for small quantities of fluids. Drops of the fluid are drawn up into a tube (1·5 mm. in diameter), so as to alternate with small drops of cane sugar of known content (Fig. 20). Water vapour will pass from the solution of which the vapour tension is the higher. By observing the edge of a drop under a microscope, it can be easily seen whether it has grown or diminished in size. If the edge of the drop remains stationary, it shows that the vapour tension and the osmotic pressure of the two fluids are equal. A series of trials is made with different strengths of sugar solution until this equality is established.

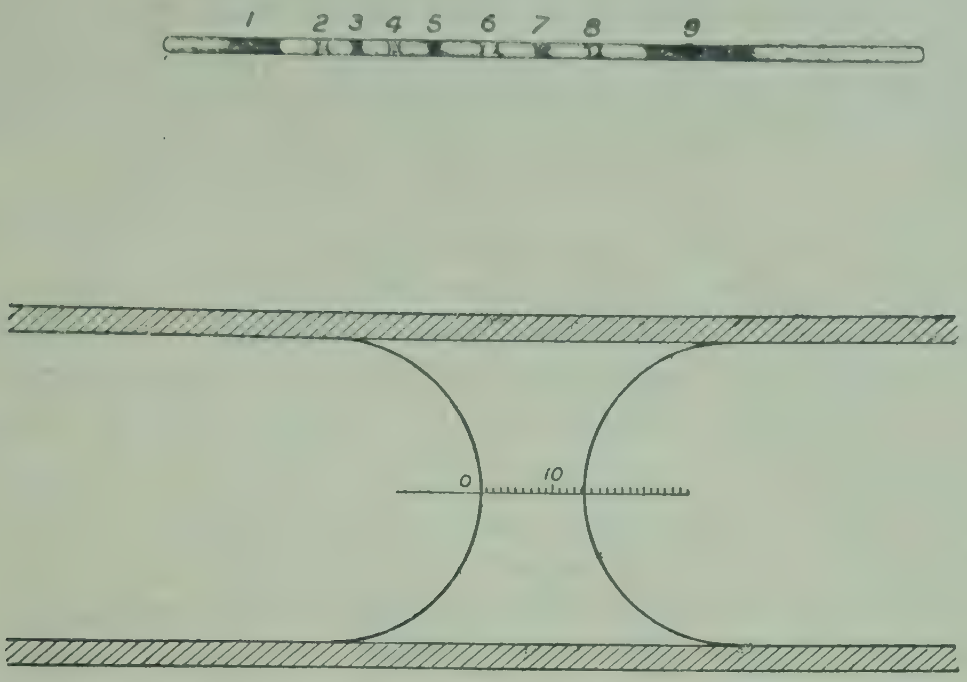


FIG. 20. Diagram to illustrate Barger's Method of Determining Osmotic Pressure. The upper figure shows the capillary tube with nine alternate drops of cane sugar and the substance under investigation.

OSMOTIC PRESSURE OF ELECTROLYTES. The conducting power of electrolytes depends on the fact that their molecules, in the act of solution in water, undergo dissociation into ions which increases with the dilution of the solution. Since each ion, as regards osmotic pressure and the related properties, acts as a whole molecule, the osmotic pressure of a *dilute* solution of NaCl is double that of an equimolecular solution of a non-electrolyte, *e.g.* glucose. Thus 0·58 per cent. NaCl is isosmotic with 3·5 per cent. glucose and not with the molecular equivalent, 1·8 per cent. If the solution is one of sodium sulphate the osmotic pressure is increased three times, since this salt dissociates into two Na ions and one SO_4 ion.

Every substance in solution evidently possesses a certain amount of potential energy in the form of osmotic pressure. This pressure is independent of the nature of the substance dissolved and is determined merely by its molecular concentration and degree of ionic dissociation. It can be used as a force for the attraction of water or, by the use of appropriate mechanisms, for the performance of mechanical work or for the production of differences of electrical potential.

"Normal" Physiological Solutions. Animal cell membranes are freely permeable to water, and hence, in conditions of equilibrium, the total osmotic concentration inside and outside the cell must always be equal. If it is temporarily disturbed, equilibrium is soon restored by passage of water towards the osmotically stronger side, by which process the cells swell or shrink. Using blood corpuscles, we may find a concentration, *e.g.* of sugar at which the corpuscles will neither increase nor diminish in bulk. This concentration varies according to the species from which the corpuscles are derived. Instead of sugar we can use sodium chloride. It is then found that a 0·65 per cent. NaCl solution is isotonic with frogs', and a 0·9 per cent. NaCl solution with mammalian red corpuscles. Since the osmotic concentration is identical in all the cells of the body and is the same as that of the blood, 0·65 per cent. and 0·9 per cent.

---

* BARGER. *Trans. Chem. Soc.*, 1904, **85**, 286.

NaCl solutions have been given the name of "*normal saline*," and are used fo maintaining tissues, respectively of frog or mammal, moist when they are unde experiment.

Since the limiting external layer of most cells seems to resemble bloo corpuscles in being impermeable to sugar and many ions, it might be thought that it was a matter of indifference what salts were used for making norma saline. This is not the case. It was shown by Ringer (Ch. XXXII) that t sodium chloride there must be added traces of calcium and potassium salts in order to obtain a saline solution suitable for keeping the heart alive and beating.

The function of a cell is thus dependent, not only on the nature of its contents and on the osmotic pressure of the surrounding fluid, but also on the nature of the ions which are in contact with its external surface. It should be noted that it is a matter of indifference what salts of calcium or potassium are used, since it is the cations themselves which are responsible for the effect. When we wish to keep a tissue alive, immersed in some solution, we employ a "Ringer's fluid" in which sodium, potassium and calcium ions are present in the proper amounts. It is customary to add also some sodium bicarbonate and sometimes sodium phosphate, so as to imitate the reaction of the normal body fluids. The accompanying table shows the composition of some commonly employed physiological saline solutions.

Physiological Saline Solutions.

	Ringer (Frog).	Ringer-Locke.	Ringer-Tyrode.	Ringer-Dale.	Buffered Ringer [C.L.E.].
Water . . .	100·0	100·0	100·0	100·0	100·0
NaCl . .	0·65	0·9	0·8	0·9	0·85
KCl . . .	0·014	0·042	0·02	0·042	0·042
$CaCl_2$ . .	0·012	0·024	0·02	0·024	0·024
$NaHCO_3$ . .	0·02	0·01–·03	0·1	0·05	—
NaH_2PO_4 . .	0·001	—	0·005	—	—
$MgCl_2$ . .	—	—	0·01	0·0005	—
Glucose . .	[0·2]	[0·1–0·25]	0·1	0·05	0·1
Na_2HPO_4 . .	—	—	—	—	0·06
$\frac{M}{1}$ H_3PO_4 . .	—	—	—	—	0·02–0·06 c.c.

For Reference

Findlay (1919). Osmotic Pressure. (London : Longmans, Green & Co.)

Lucké and McCutcheon (1932). The Living Cell as an Osmotic System and its Permeability to Water. *Physiol. Rev.*, **12**, 68.

Schreinemakers (1938). Lectures on Osmosis. (Hague : Naeff.)

NEUTRALITY AND REACTION

Most of the tissues in the human body contain about 80 per cent. of water. In every reaction taking place in the organism water plays an essential part. It is calculated that 1 litre at 22° C. contains 10^{-7} gramme ions of hydrogen, and of course the same number of hydroxyl ions. The temperature coefficient of the dissociation of water is fairly high, so that at 18° C. the hydrogen ion concentration in pure water is only $0{\cdot}78 \times 10^{-7}$, while at 37° C. it is $1{\cdot}76 \times 10^{-7}$.

When a strong acid such as hydrochloric acid is added to a strong base such as NaOH, the hydrogen and OH ions combine to form molecules of water, leaving behind only Na and Cl ions. The actual combination which takes place is thus the same whatever strong acid and strong base are used, namely, a combination of H and OH ions. In a solution containing H and OH ions, the active mass in any reaction which may take place is equal to the product of the

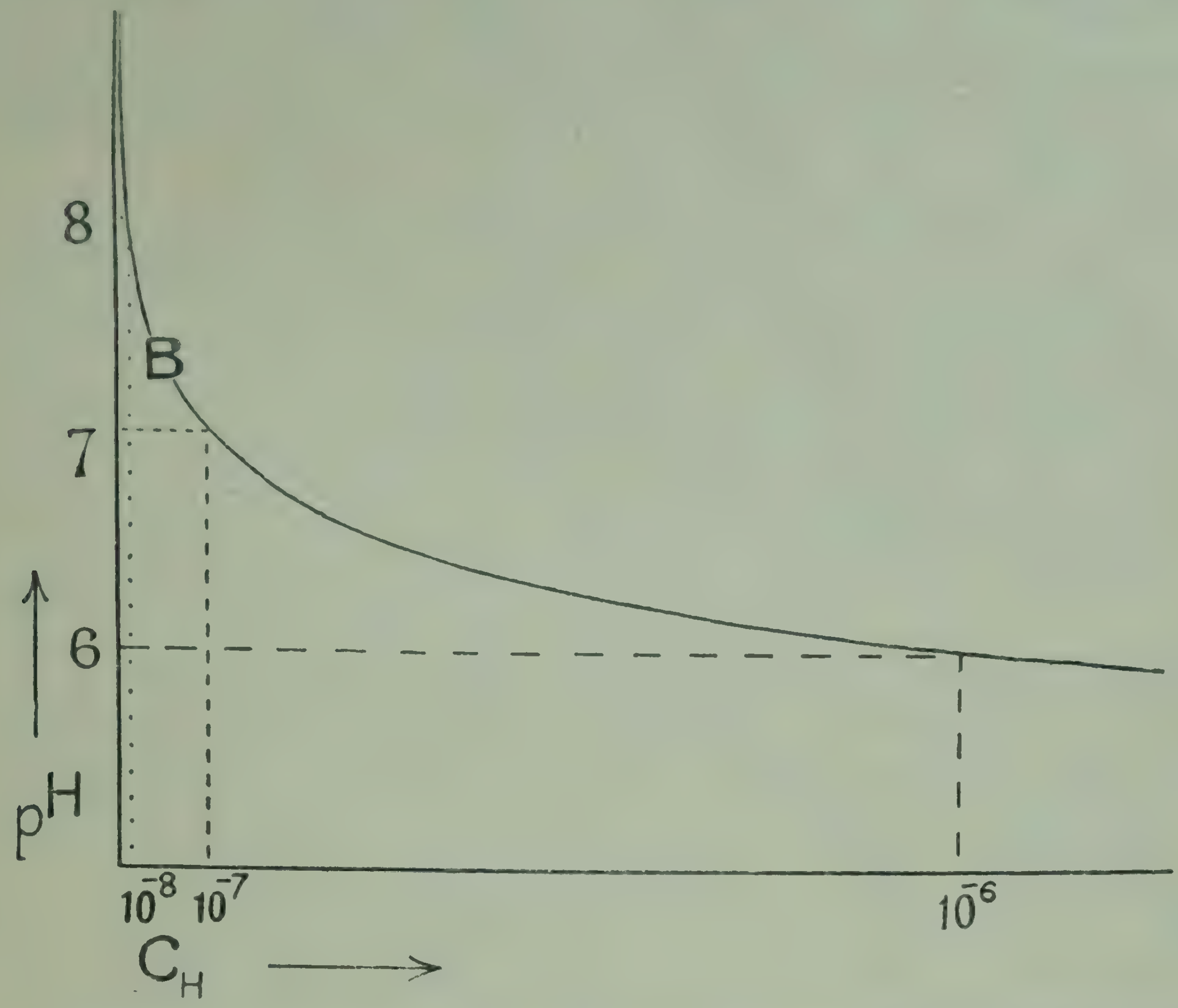


FIG. 21. Diagram to show the Relation between the two Methods of Registering Hydrogen Ion Concentration. (D. T. HARRIS.)

two concentrations H × OH. In a watery solution this is constant for any given temperature and amounts at 22° C. to 10^{-14}, increasing with the temperature; (at 37° it is $3{\cdot}1 \times 10^{-14}$). Thus the more H ions are present in a solution the fewer will be the OH ions, and the shift is considerable, since it is the product of the two and not the sum which remains constant.

The behaviour of a living cell bathed in any solution in which H ions are concerned will be determined, not by the total titratable acid or base, but by the actual concentration or activity of the H ions in the solution. It is therefore of importance to express the reaction of the body fluids in terms of H ion concentration, denoted by cH, or of H ion activity *ca*H.

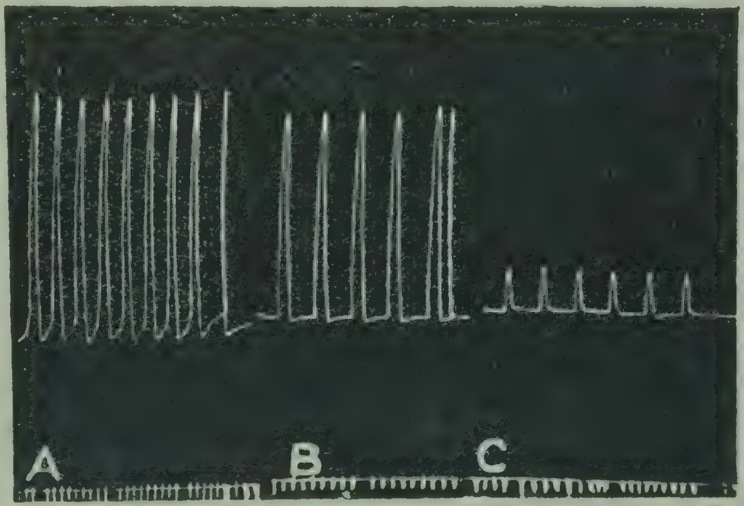


FIG. 22. Effect of Acid on the Isolated Heart of the Frog. (CLARK.)

A. Perfused with Ringer's solution with H' ion concentration of pH 7·7.

B. After perfusion for 20 minutes with faintly acid Ringer's solution, H' ion concentration, pH 6·5.

C. After perfusion of the acid solution for 80 minutes.

Owing to the inconvenience of expressions such as $1{\cdot}05 \times 10^{-7}$, $1{\cdot}3 \times 10^{-6}$, etc., it has become customary to use, not the cH figure directly, but the negative exponent of 10 as a positive number, calling it the hydrogen ion exponent or pH. Thus 5×10^{-6} is the same as $10^{-5{\cdot}3}$, and a solution having this concentration in H ions would be described as having a pH of 5·3. It must be remembered that the pH falls as the acidity increases. Moreover the change expressed by a unit difference in pH is much greater than that expressed by a unit difference in cH. Thus, while it is easy to see that the cH of 4×10^{-6} is double that of 2×10^{-6}, it is not at once obvious that a pH of 5·398 signifies an acidity which is double that of 5·699. The relation between the values of cH and of pH between cH 10^{-8} and cH10^{-6} is shown in Fig. 21.

Since H and OH ions are involved in all the chemical processes which go o in the living cell, it is not surprising that the cell functions are profoundl modified by differences in reaction. In Fig. 22 is shown the effect of a chang from $10^{-7\cdot7}$ to one of $10^{-6\cdot5}$ on the isolated heart of the frog. This change woul be produced by the addition of 0·036 milligrammes of HCl to 1 litre of distille water. A slightly higher degree of acidity, *i.e.* one of 10^{-6}, would kill th heart altogether.

“Buffer” Systems. The changes of reaction which would be caused by production of minute quantities of acid or alkali in the various chemica processes continually taking place in the cell are minimised by the presenc of certain substances in the protoplasm and fluids of the body in which the cells are bathed. These substances function as “buffers” in taking up $H^{\cdot}$ or OH' ions without a corresponding change in the reaction of the medium.

As an example of buffer action, we may consider the pH titration curves obtained by addition of NaOH to acetic acid, as compared with the curves obtained by adding NaOH to water rendered equally acid with HCl (Fig. 23). It will be seen that, whereas the addition of 0·2 c.c. of NaOH raises the pH of the water from 3·5 to 10, the pH of the acetic acid is only raised to about 3·6, and even 2·5 c.c. of alkali only raises it to about 5·5. Conversely, if we started with sodium acetate and added HCl, the pH would only slowly alter between pH 6·5 and 3·5, as compared with the alteration which would be caused by the same additions to water.

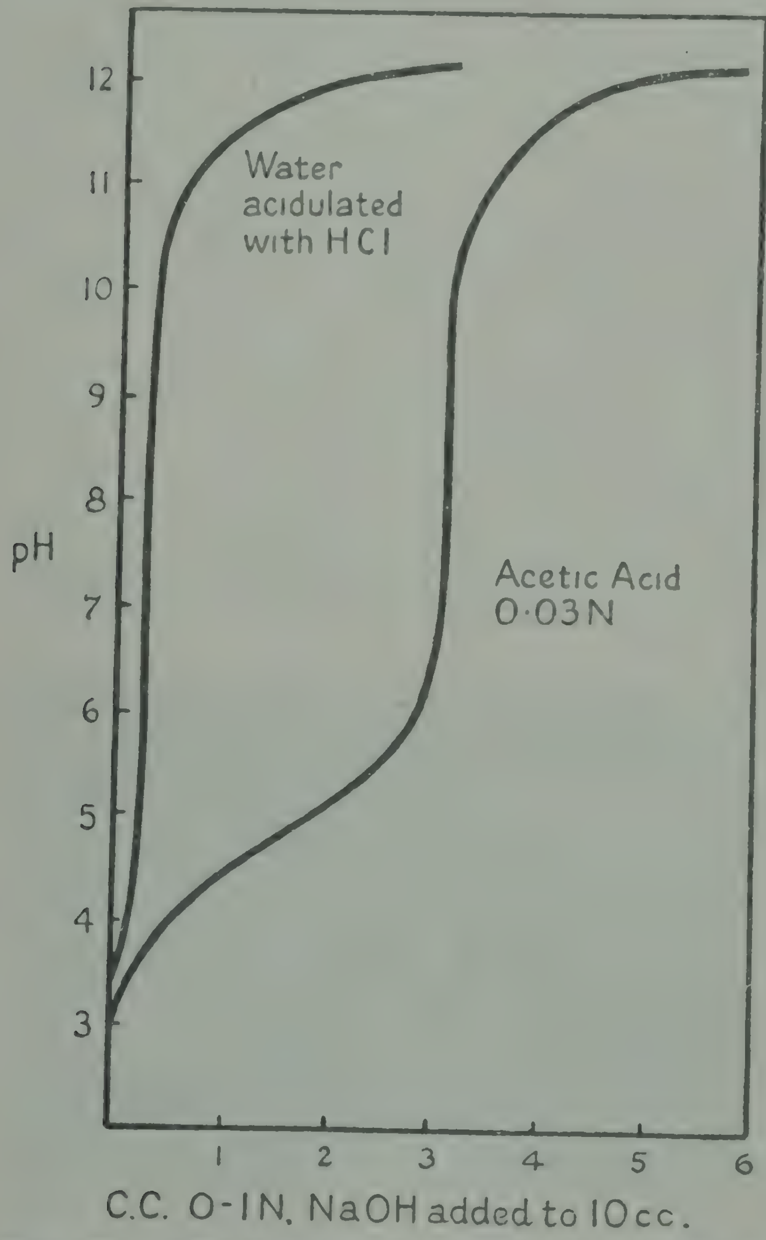


FIG. 23. pH titration curves, showing effect of adding stated amounts of 0·1N. NaOH to 10 c.c. of 0·03N. acetic acid, and to 10 c.c. water acidulated to pH 3·5 by HCl.

Two important salt buffers are the bicarbonate system and the phosphate system, the former of chief importance in the fluids of the body, the latter in the cells. The phosphate system is made up of the two phosphates NaH_2PO_4 and Na_2HPO_4. In a watery solution of NaH_2PO_4 there are present, besides the $H^{\cdot}$ and OH' ions from the water, $Na^{\cdot}$, H_2PO_4', $H^{\cdot}$, and HPO_4'' ions from the phosphate. The $H^{\cdot}$ ions are therefore present in greater quantity than in pure water, and the reaction is slightly acid. In a solution of Na_2HPO_4, the ions resulting from dissociation of the phosphate are $Na^{\cdot}$, $Na^{\cdot}$, and HPO_4''. The HPO_4'' ions combine with some of the $H^{\cdot}$ ions derived from the water to form H_2PO_4', leaving an excess of OH' ions, so that the reaction is slightly alkaline. Since it will require an equivalent weight of NaOH to convert 1 gramme molecule of NaH_2PO_4 into Na_2HPO_4, and 1 gramme molecule of HCl to effect the reverse change, a mixture of the two phosphates will absorb relatively large quantities of acid or base with but slight changes in the resulting reaction.

The acidity (*i.e.* the hydrogen ion concentration) of a mixture of the two phosphates will be proportional to $\frac{NaH_2PO_4}{Na_2HPO_4}$. When this proportion equals $\frac{1}{2\cdot5}$ the reaction will be neutral.

A still more perfect buffer system is afforded by $NaHCO_3$, when the solution is maintained in equilibrium with an atmosphere containing a constant proportion of CO_2. The solution will contain the following ions produced by dissociation.

$$\begin{aligned}
&NaHCO_3 && \rightleftharpoons Na^{\cdot} + HCO_3'.\\
&H_2CO_3 && \rightleftharpoons H^{\cdot} + HCO_3'.\\
&H_2O && \rightleftharpoons H^{\cdot} + OH'.\\
&HCO'_3 + H_2O && \rightleftharpoons H_2CO_3 + OH'.
\end{aligned}$$

H_2CO_3 is a very weak acid, so that only few hydrogen ions are produced. If an acid be added to the solution, a certain amount of the $NaHCO_3$ is decomposed with the formation of CO_2 and a neutral salt. But the CO_2 escapes from the solution, so that the only result is a diminished concentration of the $NaHCO_3$.

The cH of the solution will be proportional to $\frac{H_2CO_3}{NaHCO_3}$, and the reaction will be neutral when $\frac{H_2CO_3}{NaHCO_3} = \frac{1}{3{\cdot}75}$.

The salts of any weak acids (*e.g.* boric, acetic, etc.) or weak bases, *e.g.* ammonia, may act as buffers, and we may have complex systems in which these substances are present. Such buffer mixtures are of great value in physiological work because, by mixing two such substances in known proportions, solutions of definite hydrogen ion concentration may be prepared; such solutions only slowly change their H-ion concentrations when acids or bases are added to them. Examples of such buffer mixtures are given below.

Range of Buffer Mixtures

Mixtures of	Range of *p*H obtained.
Glycine and HCl	1·15 — 3·68
Acid K. Phthalate and HCl	2·2 — 3·8
Acid K. Phthalate and NaOH	4·0 — 6·2
Acetic acid and sodium acetate	3·24 — 6·55
$K_2HPO_4 + KH_2PO_4$	5·19 — 8·21
Boric acid, HCl and NaOH	7·8 — 10·0
Glycine + NaOH	8·58 — 12·97

Among other important buffers present in cells the proteins must be mentioned. Like the amino-acids, *e.g.* glycine shown above as a laboratory buffer, the proteins can function as either weak bases or weak acids and so absorb considerable amounts of either acid or alkali, so helping to stabilise the H-ion concentration of cells.

THE MEASUREMENT OF HYDROGEN ION CONCENTRATION

The Hydrogen Electrode. This is the classical method, and is still the standard of reference. Its use in biological material is restricted owing to the necessity for complete absence of oxygen.

The Glass Electrode. The method depends on the fact that a potential difference is set up between two solutions of different hydrogen ion concentrations, when these are separated by a very thin glass membrane. If the hydrogen ion concentration of one of the solutions is known, that of the other can be calculated from the difference of electrical potential. The method * has certain advantages, one of which is that it is not necessary to de-oxygenate the solution. It is now the most generally used method.

Indicators. In some cases the estimation can be made rapidly by means of *indicators.* These are dyes which change colour at definite values of hydrogen ion concentration. They are chiefly salts of either a very weak acid or a weak base, or the free acid or base itself, the change of colour being due to the electrolytic dissociation of the salt with the production of an ion which has a different colour from that of the undissociated acid o base. Some of the most useful indicators, and the range of their concentrations, are give in the accompanying table. In determining the hydrogen ion concentration of fluids suc as the blood, the reaction may be tested in the dialysate obtained by placing the blood in thimble of collodion, which is immersed in a small quantity of 0·85 p.c. NaCl solution,

---

* Kerridge. *Bioch. J.*, 1925, **19,** 611; *J. Sci. Insts.*, 1926, **3,** 404; Dole, 1941. The Glass Electrode. (Wiley, New York.)

COMMON INDICATORS FOR *p*H DETERMINATIONS.

Indicator.	Range, *p*H.	Acid Colour.	Alkaline Colour.
Methyl violet	0·1– 3·2	Yellow	Violet.
Thymol blue (acid range)	1·2– 2·7	Red	Pale yellow.
Bromphenol blue	3·0– 4·6	Yellow	Blue.
Methyl orange	3·2– 4·6	Red	Yellow.
Congo red	3·5– 4·7	Blue	Red.
Bromcresol green	4·0– 5·6	Yellow	Blue green.
Methyl red	4·3– 6·3	Red	Yellow.
Neutral red	6·5– 8·4	Red (blue in strong acid solutions).	Yellow.
Bromcresol purple	5·2– 6·8	Yellow	Purple.
Phenol red	6·8– 8·4	Yellow	Red.
Cresol red	7·2– 8·8	Reddish	Orange.
a-Naphthol-phthalein	7·6– 8·5	Yellow	Blue.
Thymol blue (alkaline range)	8·0– 9·6	Pale yellow	Blue.
Phenol-phthalein	8·3–10·0	Colourless	Red.
Thymol-phthalein	9·3–10·5	Colourless	Blue.
Nitrothymolsulphonephthalein	9·2–11·5	Violet	Yellow.
Alizarin yellow	10·2–12·1	Yellow	Orange.
Tropaeolin O	11·1–12·7	Yellow	Orange.

care being taken to avoid loss of carbon dioxide. The dialysate acquires the same reaction as the blood, which is determined by adding a suitable indicator (neutral red or phenol red) and matching against standard buffer solutions to which the same indicator has been added.

For Reference

MANSFELD CLARK (1928). The Determination of Hydrogen Ions. (Baltimore : Williams Wilkins.)

OGSTON (1947). The Definition and Meaning of *p*H. *Physiol. Rev.*, **27**, 228.

SURFACE ENERGY

The forces at the surface of a liquid differ from those in its interior. At the surface, any molecule, instead of being subject to the attraction of the other molecules on all sides, is exposed only to those attractions which tend to draw it inwards, so that as a result the surface is always the least possible and is in a state of tension. This *surface tension* can be measured, and is expressed in dynes per centimetre. To stretch the surface requires the expenditure of energy, while work may be done if the surface is diminished. It is surface tension which causes water to rise in capillary tubes, and the height to which the water rises may be used as one means of measuring surface tension. If, as in the case of mercury, a fluid does not moisten the capillary tube, there is a depression of the meniscus. Any substance which lowers the surface tension of water will tend to spread over its surface until it is one molecule thick. This is the case with oil dropped on water, the spreading of the oil being due to the expansion of the surface of water at the part where its surface tension is diminished by contact with oil. Wherever we have two immiscible fluids there will be a surface tension, and in the same way there will be a surface tension between liquids and solids.

The fact that protoplasm consists largely of colloids involves a heterogeneity of structure, and hence large areas of surfaces. The tension at the cell surface has been measured in various ways and found to lie usually between 1·0 and 0·1 dyne/cm. This value is very much lower than would be expected if the surface layer consisted of lipoid alone, and is explained as being due to the presence of an adsorbed layer of protein covering the lipoid membrane.[†]

The surface tension of a liquid is altered by dissolved substances. Inorganic salts cause a slight rise of surface tension of a solution in contact with air. Most organic substances

---

[*] DALE AND LOVATT EVANS. *J. Physiol.*, 1920, **54**, 167.

[†] NEWTON HARVEY and DANIELLI. *Biol. Rev.*, 1938, **13**, 319.

cause a diminution of surface tension. It was shown by Gibbs and Thomson that dissolved substances which lower surface tension tend to concentrate at the surface, on account of he fact that the free energy of the system is thereby lessened. In many cases this concentration at the surface may be so marked that the substance forms a solid pellicle (p. 15).

When we are dealing with an interface between liquid and liquid or between liquid and solid, this concentration of solute or dissolved substance at the surface is equivalent to a deposition on the other substance with which the solution is in contact. It is responsible for many cases of concentration at surfaces, for which the name *adsorption* is used (*v.* p. 74). The marvellous power of adsorption of gases by charcoal is due to the very large surface presented by this substance. In the dyeing of stuffs and the staining of sections adsorption is the chief factor involved. On chemically active solutes this process of adsorption will have a two-fold effect: that part which is adsorbed at the surface is taken out of the main solution and loses, for instance, its osmotic effect. On the other hand, the concentration of two reactive substances in the same surface may bring about their interaction, so that chemical changes are furthered which without this concentration would have gone on much more slowly.

Many surfaces such as glass, paper, or wool, in water are negatively charged. This charge affects the type of adsorption at the surface of such particles, and will tend to favour the concentration of electro-positive as against electro-negative molecules at the interface. A charge on a surface implies a diminution of the surface tension, and this is the principle of the capillary electrometer.

Since electrolytes may augment or diminish the charge at a surface, we see how important may be the constituents of a solution to the living cells with which it has to come in contact.

For Reference

Bayliss (1924). Principles of General Physiology. (Longmans.)
Bayliss (1923). Interfacial Forces and Phenomena in Physiology. (Methuen.)
Mathews (1921). Adsorption. *Physiol. Rev.*, **1**, 553.

HEAT FORMATION AND FREE ENERGY

Living cells utilise energy for a variety of purposes—for chemical synthesis, osmotic work, movement, etc. This energy is obtained, in nearly all cases, by oxidative processes, the net result of which is that foodstuffs of high potential energy content are converted into end-products of low energy content, the difference in energy content appearing as kinetic energy, one form of which is heat.

Absolute values of the internal energy of molecules are unknown: we can only measure changes, the difference appearing as heat or work. By convention it is usually expressed in calories (but 1 large calorie = 427 Kg. m.).

When a foodstuff is oxidised in the body, the amount of heat ultimately produced is the same as would be evolved if it were oxidised to the same end-products in any way, *e.g.* by burning it, outside the body. The heat evolved per mole, the difference between the potential energy of the initial substances and the final end-products, is called ΔH and is given a negative sign. Thus for the combustion of glucose:—

$$C_6H_{12}O_6 + 6O_2 = 6CO_2 + 6H_2O$$
$$\Delta H = -674 \text{ Cals.}$$

The second law of thermodynamics tells us that the cell cannot utilise heat for doing work, and if an energy change leads to work being done, the heat liberation must be proportionately less. For study of the work available, a knowledge of the change in *free energy* or ΔF is more useful than ΔH, since free energy is energy in a form capable of doing work.*

If a reaction involves a decrease in free energy (*i.e.* if ΔF is negative in sign) then that reaction *could* occur spontaneously, and could then do work, and the maximal amount of work theoretically possible would be deducible from, and n many cases approximately equal to, ΔF. It cannot be predicted that the eaction *will* occur, any more than it can be predicted that a car at the top of hill will run down—the brakes might be on! The value of ΔF is often about

---

[*] Without a knowledge of thermodynamics this must be accepted on faith. It would e pointless here to attempt to introduce the conception of entropy or other thermodynamical theory.

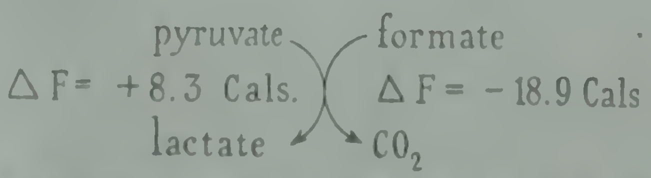
the same as ΔH, but not necessarily so. Thus for combustion of glucose, $\Delta F = -686$ Cals. Many synthetic processes that occur in living cells have a large *positive* ΔF, *i.e.* they involve *absorption* of energy and *cannot* be spontaneous reactions; just as a car at the bottom of a hill cannot get to the top without energy being expended on it. Such a reaction can only occur if coupled thermodynamically with another reaction of suitably large negative ΔF. For example, lactic acid can be oxidised to pyruvic acid with liberation of energy :—

$$\text{CH}_3\text{CHOH.COO}' \longrightarrow \text{CH}_3\text{CO.COO}' + 2\text{H}^{\cdot}$$
$$\Delta F = -8{\cdot}3 \text{ Cals.}$$

But the reverse reaction cannot occur spontaneously since ΔF would be positive. But by coupling it with a reaction that has a large negative ΔF, viz. oxidation of formate to CO_2 :—

$$\text{COOH}' + \text{H}_2\text{O} = \text{HCO}_3' + 2\text{H}^{\cdot}$$
$$\Delta F = -18{\cdot}9 \text{ Cals.}$$

the enzymes of *B. coli* can simultaneously reduce pyruvate to lactate in the coupled reaction :—



with a net change in free energy of $(-18{\cdot}9 + 8{\cdot}3) = -10{\cdot}6$ Cals. which is evolved as heat. This may be compared with a funicular railway, in which a heavy car at the top can lift a light one at the bottom, if the two are linked by a cable.

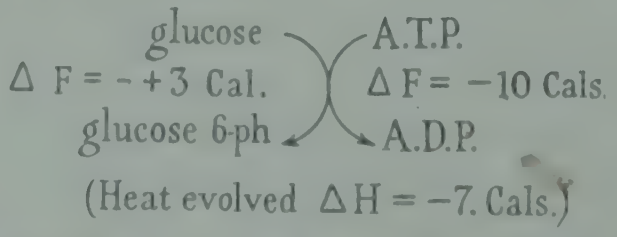
Energy-rich bonds. One of the most important substances in nature is a compound of adenosine with phosphoric acid (*v.* p. 52). It probably plays a part in all biological synthesis involving large energy changes as well as in disintegrative changes such as glycolysis and oxidation. Adenosine triphosphate (A.T.P.) may be represented diagrammatically,

$$\textcircled{\text{Ad}} \underset{(i)}{-} \text{HPO}_3 \underset{(ii)}{\sim} \text{HPO}_3 \underset{(iii)}{\sim} \text{H}_2\text{PO}_3$$

The phosphate groups can be split off one at a time and those at (iii) and (ii) carry with them about 10 Cals. of energy per mole, which can be transferred without apparent loss to other compounds.* Scission at bond (i) only carries about 2 Cal. These energy-rich groups are often designated $\sim$ Ph, and the one most frequently exchanged is the terminal one, with formation of adenosine diphosphate (A.D.P.), or the reverse.

$$\underset{\text{(adenosine triphosphate)}}{\textcircled{\text{Ad}} - \text{HPO}_3 \sim \text{Ph} \sim \text{Ph}} \rightleftharpoons \underset{\text{(adenosine diphosphate)}}{\textcircled{\text{Ad}} - \text{HPO}_3 \sim \text{Ph}}$$
$$\Delta F = -10 \text{ Cals.}$$

For example, glucose can by the appropriate enzyme be converted into glucose 6-phosphate in the coupled reaction :—



Further examples of these important reactions are given in Chapter X.

---

* Lippmann. *Adv. in Enzymology*, 1941, **1**, 99.

CHAPTER VI

THE PROPERTIES OF COLLOIDS

The term colloid, from κσλλη, glue, was first introduced in 1861 by Thomas Graham, Professor of Chemistry at University College. Graham divided all substances into two classes, viz. *crystalloids*, including such substances as salt, sugar, urea, which could be crystallised with ease, diffused rapidly through water and were capable of diffusing through dead animal membranes; and *colloids*, which included substances such as gelatin, gum, egg-albumen, starch and dextrin, were non-crystallisable, formed gummy masses when their solutions were evaporated to dryness, diffused with extreme slowness through water and would not pass through animal membranes. The process of *dialysis* was therefore introduced by Graham for the separation of crystalloids from colloids. Although the broad distinction drawn by Graham between colloids and crystalloids still holds good, some of the criteria by which he distinguished the two classes are no longer strictly applicable. For instance, it has been shown that many typical colloids, such as proteins, can be obtained in a crystalline form. On the other hand all gradations exist between substances, such as egg-albumin, which are practically indiffusible, and those, such as common salt, which are very diffusible. Colloids exist under two conditions:

(1) In a state of solution or pseudo-solution, in which they form *sols*, and are distinguished as hydrosols when the solvent is water; and

(2) In a solid state, in which a relatively small amount of the colloid sets with a large amount of a fluid, such as water, to form a jelly. This solid form, *e.g.* gelatine, is known as a *gel*. Such a watery jelly is known as a *hydrogel*.

A number of these colloidal substances can be shown on purely chemical grounds to consist of monstrous molecules. Thus the molecular weight of hæmoglobin is at least 64,000, and one must ascribe similar high molecular weights to such substances as egg-albumin and globulin. Still greater must be the molecular size of such substances as the cell proteins, which may form a gigantic complex, to which it would probably not be an exaggeration to ascribe a molecular weight of over 100,000. Chemical complexity is not, however, a necessary condition of the colloidal state, as is shown by the existence of colloidal silica, and, indeed, of colloidal solutions of practically any given substance, including many elements.

All so-called colloidal solutions are in reality suspensions of particles, the so-called *disperse phase*, in a fluid or *continuous* phase. The fluid may be any liquid or solution whatever. Thus, milk is a suspension of oil droplets in a watery phase, and butter is a colloid consisting of water suspended in a continuous phase of oil. The conversion of milk into butter is an example of phase reversal.

Of special interest are the hydrosols of the metals. Faraday long ago pointed out that, on treating a weak solution of gold chloride with phosphorus, it underwent reduction with the formation of metallic gold. The gold was not precipitated, but remained in suspension or pseudo-solution, giving a deep red * or a blue liquid, according to the conditions under which the reaction was affected. This solution was homogeneous in that it could be filtered without change and could be kept for months without deposition of the gold. The latter was, however, thrown down on addition of mere traces of impurity, though greater stability could be conferred on the solution by adding to it a little gelatin. In 1899 Bredig showed how similar hydrosols might be prepared from a number of different metals, viz. by the passage of a small arc or electric sparks between metallic terminals submerged in distilled water. If for example the terminals be of platinum, the passage of the current is seen to be accompanied by the giving off of brown clouds, which spread into the surrounding fluid.

---

* Ruby glass is a colloidal "solid" solution of gold in a mixture of silicates.

These clouds consist of particles of platinum of all sizes; the smaller—which are ultr microscopic in size, *i.e.* from 5 m.μ to 40 m.μ—remain in suspension, and we obtain a brow fluid which can be filtered through paper or even through a Berkefeld filter without losing i colour. It may be kept for months without any deposit taking place. The addition minute traces of electrolytes precipitates the platinum particles, leaving a colourless flui

Colloidal solutions or sols may be divided into two classes, *emulsoids* an *suspensoids*, according as they may be regarded as suspensions of liquid, or a suspensions of solid particles, in liquid.

Most protein solutions are emulsoids, while the metallic sols are suspensoid Dilute egg-white is an emulsoid, but if it be boiled, although no visible precipitation is produced, the fine particles are coagulated and it behaves as a suspensoid

Properties of Gels. The firm mass to which a solution of gelatin sets o cooling is a typical hydrogel. It is clear, hyaline, apparently structureless an elastic. It may be regarded as formed by the separation of the warm pseudo solution of gelatin into two phases: first, a solid phase, rich in gelatin an forming a tissue or meshwork in the interstices of which is embedded th second phase, consisting of a very weak solution of gelatin.

In many cases the more fluid part of the gel is practically pure water. I a dry atmosphere the gel loses water and becomes shrivelled and dry, but i some cases, *e.g.* gelatin, it can resume its former size and characters on immersion in water. Other gels, such as silicic acid or ferric hydrate, lose the power o swelling up after drying. The change in them is therefore irreversible. A ge adheres to the last traces of water with extreme tenacity. In consequence o its structure, it presents an enormous extent of surface on which adsorption ca take place. At this surface the vapour tension of fluids is diminished, as wel as the osmotic pressure of dissolved substances.

Many colloids, even those such as starch or gelatin which are insoluble i cold water, absorb water and exhibit a phenomenon, viz. "*Quellung*" o imbibition, which in some cases it is impossible to distinguish from the process of solution. When soaked in aqueous solutions, the amount of water taken up varies inversely with the osmotic pressure of the solution. In many respects the insoluble emulsoid colloids when soaked in aqueous solutions, behave as though they were osmotic cells with boundaries capable of expansion. Indeed, we may regard them as such cells filled with colloid to which their surfaces are impermeable. When dry, colloids can exert a considerable force when caused to swell up by moistening.

The huge aggregates of molecules, which distinguish the colloidal state, form a system with considerable inertia, so that there is a tendency to the establishment of conditions of false equilibrium. Once a state is established, it is necessary to overstep widely the conditions of its formation in order to destroy it. Thus a 10 per cent. gelatin solution sets at 21° C. but does not melt until warmed to 29·6° C. Solutions of agar in water set at about 35° C. but do not melt under 90° C. A gel of gelatin takes twenty-four hours after setting to attain a constant melting point. This phenomenon is called *hysteresis*.

On account of the extent of surface it is practically impossible to wash out the inorganic constituents from a gel. The diminution of the osmotic pressure of many dissolved substances at surfaces causes the concentration at the surface of the solid phase to be greater than that in the surrounding medium. Thus, if dry gelatin be immersed in a salt solution it will swell up, but the solution which it absorbs will be more concentrated than the solution in which it is immersed, so that the proportion of salt in the latter will be diminished. When, however, equilibrium is established between a gel and the surrounding fluid, it is found to present no appreciable resistance to the passage of dissolved crystalloids. Thus salt or sugar diffuses through a column of solid gelatin as if the latter were pure water. But gels are impermeable to other colloids in solution. This impermeability is made use of in the separation of crystalloids from colloids by dialysis, membranes used in this process being generally irreversible gels (*e.g.* vegetable parchment, animal membranes, collodion or cellophane). Othe gels, such as tannate of gelatin or copper ferrocyanide, are not only impermeable to colloids, but also to many crystalloid substances. These membranes can therefore be used for the determination of the osmotic pressure of those crystalloids.

Properties of Hydrosols. Hydrosols of substances such as dextrin or egg-albumin may be made of almost any concentration. If a solution of egg-albumin be concentrated at a low temperature, it becomes more and more viscous and finally solid. But there is no distinct point at which the fluid passes into the solid condition. Much discussion has arisen whether hydrosols are to be regarded as true solutions or as suspensions. The homogeneity or heterogeneity of a solution is clearly a matter of degree. In a true solution the molecules of the solute are equally diffused throughout the molecules of the solvent, and it is impossible, without the application of energy, to separate one from the other. Thus filtration or sedimentation leaves the composition of the solution unchanged. It is true that by the employment of the semi-permeable membrane, it is possible to separate solute from solvent, but in this case the force required to effect the filtration may be enormous. The measure of the force required is the osmotic pressure of the solution.

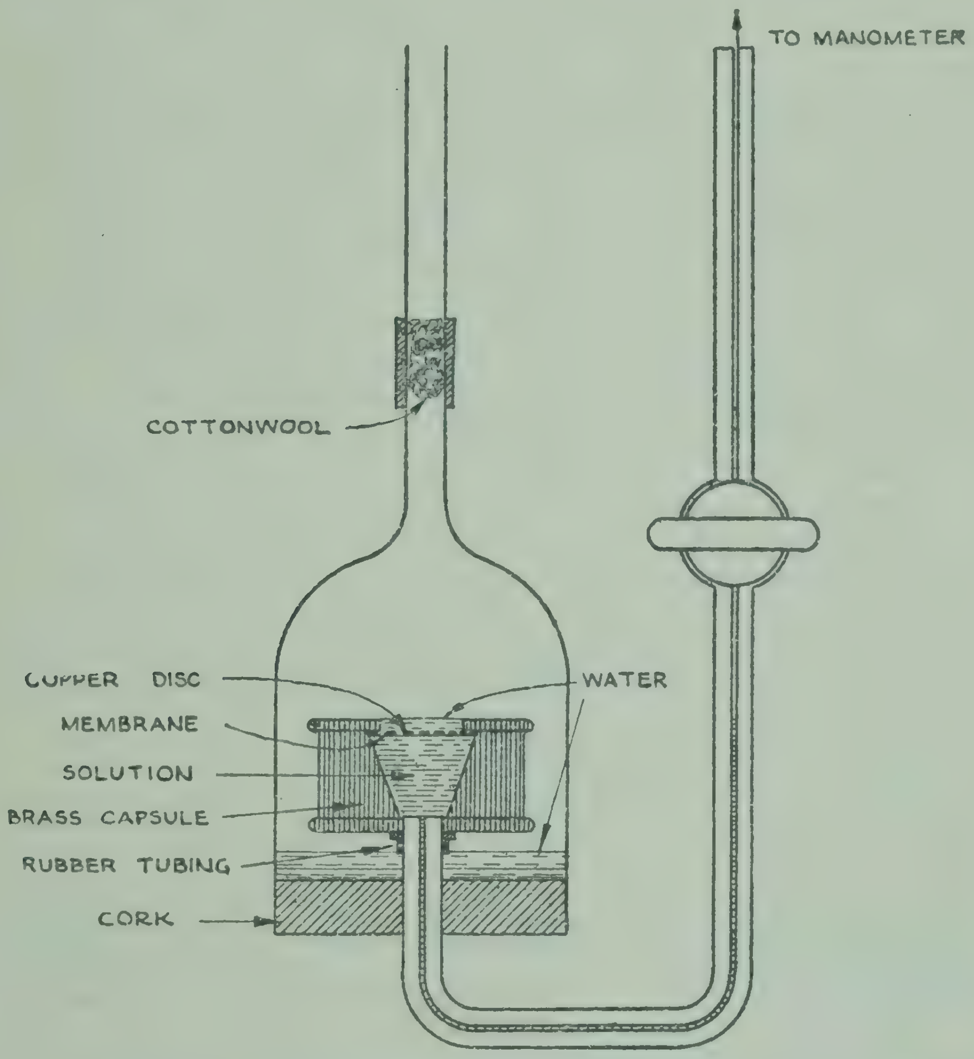


Fig. 24. Govaerts' Osmometer (Verney's modification). (Kerridge.)

Osmotic Pressure. Starling * showed that it is possible to determine the osmotic pressure of colloidal solutions directly, taking advantage of the fact that colloidal membranes, while permitting the passage of water and salts, are impermeable to colloids in solution. In this way it was found that blood serum, containing 7 to 8 per cent. of proteins, had an osmotic pressure of 25 to 30 mm. Hg, which would correspond to a molecular weight of this protein, of about 30,000.

A convenient osmometer for this purpose is illustrated in Fig. 24. The membrane is of cellophane. The serum or other solution is placed in the brass capsule, and water on the upper surface of the membrane. The level of serum in the glass capillary is kept constant by applying counter-pressure from a manometer, to balance the osmotic pressure, and this counter pressure is read off when equilibrium is reached.†

* Starling. *J. Physiol.*, 1895, **19**, 323; 1899, **24**, 317.

† A simpler form is described by Keys and Taylor. *J. Biol. Chem.*, 1935, **109**, 47.

The existence of an osmotic pressure in colloidal solutions was established by Moor and Roaf[*] in the case of hæmoglobins, proteins and soaps, and by Bayliss in the case c colloidal dyes.[†] The osmotic pressure of hæmoglobin has been found by Adair[‡] to correspond to a molecular weight of about 68,000, *i.e.* about 4 times as great as the minima molecular weight deduced from its combining power with oxygen. The osmotic pressur is often smaller than would be expected from the molecular weight of the substance, owin to molecular aggregation. Thus the aggregate of colloidal silica must be many, probabl thousands of, times larger than the molecule as represented by H_2SiO_3. The osmoti pressure being proportional to the number of separate molecules or groups of molecules in given volume of solution, to which a given membrane is impermeable, the larger the aggregate the smaller would be the osmotic pressure of the solution.

It is in consequence of the huge size of the molecular aggregates that colloida solutions, such as starch or glycogen, display such a trivial osmotic pressure We can divide colloidal solutions into two classes, viz. those which almost form true solutions and present a feeble osmotic pressure, and those which form suspensions and therefore exert no osmotic pressure. In inorganic colloids such as arsenious sulphide, Picton and Linder have shown that all grades exis between true solutions and suspensions. With increasing aggregation of the molecules, the suspension becomes coarser and coarser until finally the sulphide precipitates.

Heterogeneity. We have every reason to believe that the very complex molecules forming the cell proteins may have molecular weights even greater than 30,000. When, however, we arrive at molecular weights of these dimensions, the disproportion between the size of the molecules and those of the water becomes so great that a homogeneous distribution of the two substances, solute and solvent, is no longer possible. The size of a molecule of water has been reckoned to be $0{\cdot}7 \times 10^{-8}$ mm. A molecule 10,000 times as large would have a diameter of $0{\cdot}7 \times 10^{-4} = 0{\cdot}07\mu$, a size just within the limits of microscopic vision. Long before molecules attained such a size they would no longer react according to the laws of solutes, but would possess the properties of matter in mass. They have a surface of measurable extent, and their relations to the molecules of solvent will be determined by the laws of adsorption at surfaces. As a matter of fact we find that such solutions present an amazing mixture of properties, some of which betray them as mechanical suspensions, while others show them to be capable of participating in chemical reactions such as those studied in the simpler compounds in solution usually dealt with by the chemist.

THE ULTRA-CENTRIFUGE. It has been found by Svedberg[§] that protein solutions and other colloids show partial sedimentation when submitted to extremely rapid centrifugalisation, and he has calculated the apparent molecular weights of various proteins from their observed speeds of deposition. It would appear from his researches that most proteins have molecular weights of 17,600 or approximate multiples of this, *e.g.*:

Egg albumin	43,800
Hæmoglobin	69,000
Serum albumin	70,200
Serum globulin	167,000
Nucleic acid	15,000 to 1,000,000
H-Hæmocyanin	6,630,000

Optical Behaviour of Hydrosols. Nearly all colloidal solutions present what is known as the Faraday-Tyndall phenomenon. When a beam of light is passed through an optically homogeneous fluid, the course of the beam is invisible. A beam of sunlight falling into a dark room is rendered visible by impinging on and illuminating the dust particles in its course. Each of these particles, being illuminated, acts as a centre of disperson of the light, so that the course of the beam is apparent to a person standing on one side of it. Tyndall showed that, if the particles were sufficiently minute, the light dispersed by them at right angles to the beam was polarised. This can be easily tested by looking at the beam through a Nicol's prism. The Tyndall phenomenon may be regarded as a test for the presence of ultra-microscopic particles, varying in size from 5 to 50 mμ. The phenomenon is perhaps too sensitive to be taken as a proof that a fluid presenting it is a suspension rather than a solution. It is shown, for instance, by solutions of many bodies of high molecular weight, such as raffinose (a trisaccharide) or the alkaloid brucine (Bayliss).

A particle having a diameter less than half the wave-length of light, *i.e.* about 300 λ o 0·3 μ, cannot be clearly distinguished under any power of the microscope. The fact that a ultra-microscopic particle may serve as a centre for dispersal of light may be used fo endering the *position* of such particles visible under the microscope. For this purpose a trong beam of light is passed in the plane of the stage of the microscope through a cell ontaining the hydrosol, which is then examined under a high power. On examining a ilute gold sol with this apparatus, we see a swarm of dancing points of light, "like gnats in he sunlight," which move rapidly in all directions, rendering it almost impossible to count heir number in the field. The coarser particles present slight oscillations similar to those ong known as the Brownian movement. The smallest demonstrable particles show a ombined movement, consisting of a translatory movement, in which the particle passes apidly across the field in one-sixth to one-eighth of a second, with one of oscillation of much horter period. The representation of the course of such a particle is given in Fig. 25.

---

[*] MOORE AND ROAF. *Bioch. J.*, 1907, **2**, 34.

[†] BAYLISS, W. M. *Proc. Roy. Soc.*, 1909, **81B**, 269; 1911, **84B**, 229.

[‡] ADAIR. *Proc. Roy. Soc.*, 1925, **109A**, 292.

[§] SVEDBERG. *Transac. Faraday Soc.*, 1930, **26**, 740. *Nature*, 1937, **139**, 1051; *Pro R. S.*, 1939, **127B**, 1.

The size of the smallest particles revealed in this way may amount to 0·005 μ. Not all colloidal solutions show these particles in the ultra-microscope. In some cases this is due simply to the small size of the particles, and the addition of any substance, which causes aggregation and therefore increase in the size of the particles, will bring them into view. In others the absence of optical heterogeneity may be due to the coincidence of the refractive indices of the two phases of the hydrosol, or to the absence of any clear dividing surfaces between the two phases. In any case it is not the particle itself which is thus rendered visible, but only the light scattered from it.

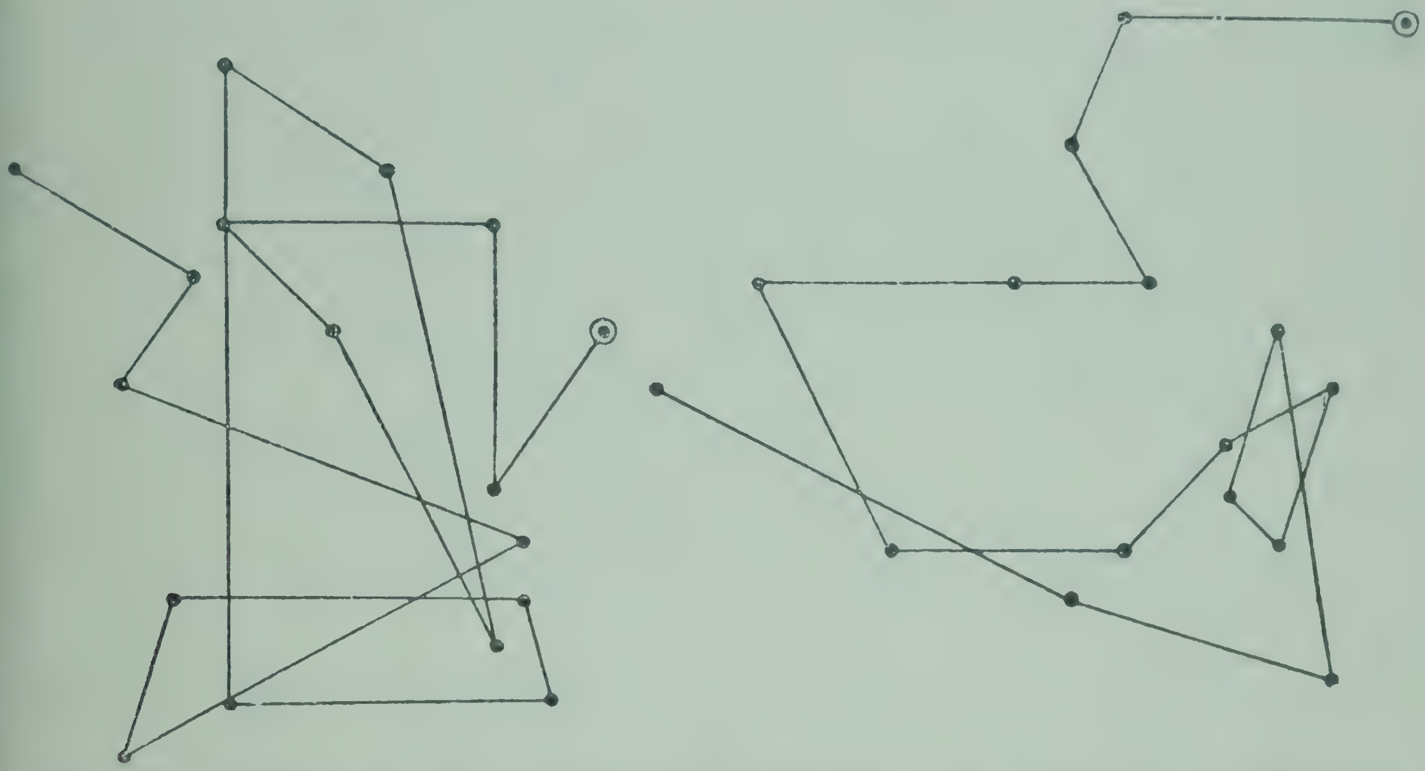


FIG. 25. Movements of two Particles of India-rubber Latex in Colloidal Solution, recorded by cinematograph and ultra-microscope. (HENRI.)

ELECTRICAL PROPERTIES OF COLLOIDS

In the case of many hydrosols the ultra-microscopic particles of which they are composed carry an electric charge which, according to the nature of the solution, may be either positive or negative. On this account the particles move if placed in an electric field, and the direction of their movement reveals the nature of their charge. Thus colloidal ferric hydrate is electro-positive and travels to the cathode. The presence of a charge is not, however, a necessary condition for the stability of a colloidal solution. Thus the proteins of serum globulin in a neutral saline solution, or of gelatin, present no drift when exposed to a strong electric field.* In certain colloids, the proteins for instance, the charge varies according to the conditions under which they are brought into solution. At a particular hydrogen ion concentration, differing for different proteins, and called the *isoelectric point*, the protein is uncharged and does not migrate in an electric field. Moreover, at the isoelectric point, proteins show minimal solubility, viscosity, osmotic pressure, etc. Advantage is often taken of this to separate or purify proteins by precipitation at the isoelectric point, or *isoelectric precipitation*. When in solution on the acid side of their isoelectric points protein particles carry a positive charge, whereas on the alkaline side (and at most physiological pH values) their charge is negative.

---

* HARDY. *J. Physiol.*, **1905, 33, 251.**

The iso-electric pH of some important proteins are given below :—

Edestin	6·9
Reduced hæmoglobin	6·8
Oxyhæmoglobin	6·6
Serum globulin	5·4
Egg albumin	4·8
Serum albumin Casein Gelatin	4·7

The charged condition of these particles must play a considerable part in keeping them asunder and therefore in preventing their aggregation and precipitation. This is shown by the fact that any agency which will tend to discharge them will often cause precipitation and coagulation. Among such agencies is the passage of a constant current and the coagulative effects of neutral salts. Thus, any of the colloids we have mentioned, ferric hydrate, silica, gold or boiled albumin, are thrown down by the addition of traces of neutral salts, and it is found that in this process they carry with them some of the ion with the opposite charge to that of the colloidal particle. Thus in the precipitation of the electro-positive ferric hydrate, the negative ion of the salt is the determining factor, the coagulative power increasing rapidly with the valency of the acid. On the other hand, in the precipitation of a gold sol, the electro-positive ion is the effective agent, and here again the coagulative effect is enormously increased by a rise in valency. This is shown by the fact that in the coagulation of gold, barium chloride with the divalent $Ba^{\cdot\cdot}$ is 7 times as powerful as K_2SO_4 containing the univalent $K^{\cdot}$; whereas, in the precipitation of the electro-positive ferric hydrate, K_2SO_4 with a divalent SO_4'' is 400 times as effective as $BaCl_2$.

Electrophoresis. The mobility of protein molecules in an electric field, *i.e.* the velocity with which the protein molecules move (in cm./sec. per unit gradient in potential, *i.e.* per volt/cm.) can be measured by putting a known potential between electrodes, at a measured distance apart, dipping into the solution of protein, and then, after a measured interval of time, determining the concentration of protein at various distances from the electrodes. By this method Tiselius has shown that mixtures of proteins can be separated, without chemical interference.

Other Forces. In addition to the electrical forces related to ionisation, there are other less well understood forces that may play a part in determining colloidal behaviour.

Long-range colloidal forces, probably electrical, but of unknown origin, can operate over distances of the order of 500 A. Enzymes can act when separated from their substrates by films 100 A. or more in thickness, perhaps due to such forces.

Hydrogen bonds can form connections with other atoms, *e.g.* N or O, and such bonds certainly occur in proteins.

The Hofmeister Series. The action of ions on emulsoid colloids, *e.g.* proteins, is somewhat more complex than that described above, since both positive and negative ions exert an influence which is dependent on the nature of the ions. When gelatine is immersed in aqueous solutions, for example, its swelling is hastened by NaCl, NaBr, KCl, NH_4Cl and $NaNO_3$ and retarded by sodium sulphate, tartrate, citrate and acetate. It is possible to arrange certain ions in the order of their efficacies in affecting this and certain other properties of emulsoid colloids. This series, called the *Hofmeister series*, runs :—

$$\text{Sulphate, Tartrate} < \text{Acetate} < \text{Cl} < \text{NO}_3 < \text{I} < \text{SCN}$$

which exert a tendency, in the order shown, to precipitate positively charged (*i.e.* acidified) protein, while the cations

$$\text{Li} > \text{Na} > \text{K} > \text{NH}_4$$

favour precipitation of positively charged protein in the order shown. In alkaline solutions the order of each series is reversed. According to Loeb, if attention is given to hydrogen ion concentration, the series becomes altered and is then simply one of valencies. But it is probably important that the Hofmeister series also holds good for many aspects of the physiological effects of various ions in the tissues.

Adsorption. The charged condition of a colloidal particle makes it behave in an electric field in much the same way as a charged ion of an electrolyte, and this similarity extends also to its chemical behaviour, so that we can have a class of combinations formed resembling in many respects chemical compounds, but differing from them in the absence of definite quantitative relations between the reacting substances. This class of compounds of varying composition has been designated *adsorption complexes.*

The factors involved in the formation of adsorption combinations are :

(1) Extent of surface. In a colloidal solution this is enormous in proportion to the mass of substance in solution. Thus a 10 c.c. sphere with a surface of 22 sq. cm., if reduced to a fine powder consisting of spherules of 2·5 $m\mu$ in diameter, will have a surface of 20,000,000 sq. cm., *i.e.* nearly half an acre. At the whole of this surface adsorption may take place, involving the concentration of dissolved electrolytes, ions or gases.

(2) Chemical nature of particle, residual valencies, etc.

(3) Electric charge and other forces on the surface. The sign of this may be determined by the chemical nature of the colloid and its relation to the electrolytes in the surrounding medium.

The conditions which determine the formation of these adsorption complexes can be studied in their simplest form on the adsorption of dyestuffs by substances such as paper. If we take a series of solutions of a dye, such as Congo red, in progressively diminishing concentration, and place in each solution the same amount of filter-paper, we find that a part of the dye is taken up by the paper, and the proportion taken up is larger the more dilute the solution. This relation is in accordance with what Bayliss called the *law of adsorption.** This is illustrated by the following Table :

Concentration of Solution.		Proportion of Dye in Solution.	Proportion of Dye in Paper.
Initial.	Final.	Per cent.	Per cent.
0·014	0·0056	40	60
0·012	0·0024	20	80
0·010	0·0009	9·3	90·7
0·008	0·0003	4	96
0·006	0·00008	1·3	98·7
0·004	—	trace	practically all
0·002	—	trace	practically all

If put into the form of a curve, where the ordinates represent the percentage of dye left in solution and the abscissæ the original concentration of the solution, the curve only approaches the axis (*i.e.* zero concentration) asymptotically. In other words, the more dilute the original solution the more firmly the dye is adsorbed. A consequence of this is that when we attempt to wash out an adsorbed substance by the pure solvent, the removal (or *elution*) gets slower and slower. Elution can be hastened by the use of another solvent, of a nature that reduces adsorption. Similar relations are found to exist between proteins and electrolytes. By continuously washing a protein, *e.g.* gelatin, with distilled water, the removal of electrolytes becomes slower and slower, and it is impossible within finite time to get rid in this way of the last traces of ash.

CHROMATOGRAPHY.† An interesting application of the phenomena of adsorption is the technique of chromatography, first introduced by Tswett, by which dissolved substances can be separated from one another. A column of a powdered adsorbent is packed into a vertical glass tube, and the solution to be analysed poured on and allowed to percolate ; if the components to be analysed are coloured (*e.g.* plant pigments dissolved in petroleum ether) and the adsorbent white (*e.g.* calcium carbonate) it is seen that different coloured layers

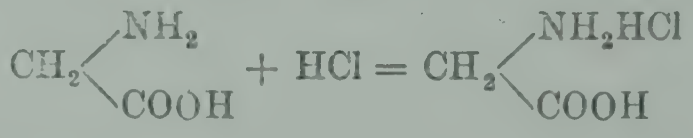
---

* **BAYLISS, W. M.** *Bioch. J.*, **1906, 1, 175.**

† WILLIAMS. An Introduction to Chromatography. 1946. (Blackie & Son Ltd.)

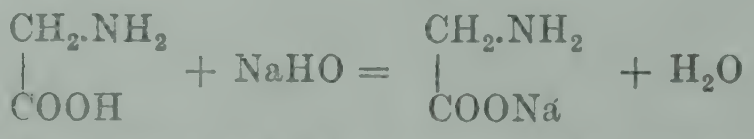
are formed owing to varying intensity of adsorption of different pigments. I
now the solution is followed through by pure solvent (*e.g.* petroleum ether) th
different bands move further apart and are separated by white intervals. Th
coloured zones can now be removed separately and the pigments eluted (*e.g.* b
alcohol). Modifications of this method allowing of the separation also o
colourless substances from one another have been developed, and have prove
to be of great value.

Colloidal Ions. Although the chemical behaviour of colloids is largel
determined by surface phenomena, it is also conditioned by the chemical structur
of the colloid molecule and by the charge carried by the latter. A good exampl
of these combinations is presented by globulin, the behaviour of which has bee
studied by Hardy.* Four states can be recognised both in the solid conditio
and in solution, viz. globulin itself, compounds with acid or with alkali, an
compounds with neutral salt. The amount of acid and alkali combining wit
the globulin is indeterminate, the effect of adding either acid or alkali to th
neutral globulin being to cause a gradual conversion of an opaque, milk
suspension into a transparent solution. On drying HCl-globulin, the solid is
found to contain all the chlorine used to dissolve it. The acid may therefore b
regarded as being in true combination. Both acid and alkali globulins act as
electrolytes, the globulin being electrically charged and taking part in the
transport of electricity. It is, in fact, present as an ion, charged positively or
negatively, according to the reaction. In order to produce the same degree of
solution, the concentration of the alkali added must be double that of the acid.
The relation of globulin to acids and alkalies is similar to that of the so-called
ampholytes or amphoteric substances, such as the amino-acids. An amino-acid
such as glycine can react as a basic anhydride with other acids thus :



$$CH_2\begin{cases}NH_2\\COOH\end{cases} + HCl = CH_2\begin{cases}NH_2HCl\\COOH\end{cases}$$

or as an acid with bases :



$$\begin{matrix}CH_2.NH_2\\|\\COOH\end{matrix} + NaHO = \begin{matrix}CH_2.NH_2\\|\\COONa\end{matrix} + H_2O$$

Like these, too, globulin forms soluble compounds with neutral salts. An
amphoteric electrolyte thus acts as a base in the presence of a strong acid and
as an acid in the presence of a strong base.

According to later developments of physical chemistry, amino-acids more probably have what is called a *zwitterion* structure : $^+NH_3.R.COO^-$. If this is so, then the action of acids and alkalies differs from that previously accepted, for the zwitterion forms a "dipole," positive at one end and negative at the other. Addition of a strong acid therefore acts, not by enhancing the basic property of the NH_2 group, so causing the molecule to acquire a positive charge, but by depressing the ionisation of the carboxyl, *i.e.* the negatively charged group :

$$^+NH_3.R.COO^- + H^+Cl^- = {}^+NH_3.R.COOH + Cl^-$$

The zwitterion theory can also be applied to polypeptides and proteins.†

From true electrolytes, colloidal solutions differ in the fact that their particles are of varying size according to the conditions in which they exist, and carry varying charges of electricity, whereas an ion such as $Na^{\cdot}$ or Cl' has a mass which is constant for the ion in question and always carries the same electric charge. The charged particles of an acid- or alkali-globulin may be distinguished therefore as *colloidal* or *pseudo-ions*. Whether a protein has a positive or a negative charge depends chiefly on the hydrogen ion concentration of the medium. At the *isoelectric point* it carries no charge (or, according to the zwitterion theory, there is complete ionisation of both groups).

---

* Hardy. *J. Physiol.*, 1905, **33**, 251.

† Birch and Harris. *Bioch. Journ.*, 1930, **24**, 1080 ; **Richardson.** *Proc. Roy. Soc.*, 1934, **115B**, 121.

COMBINATIONS BETWEEN COLLOIDS. Besides the compounds between colloids and electrolytes, combination, or at least interaction, takes place between different colloids. Many colloids are precipitated by other colloidal solutions. This effect is always found to occur when the colloidal solutions carry opposite charges. Thus ferric hydrate in colloidal solution is precipitated by colloidal silica or colloidal gold, both colloids being thrown out of solution. On the other hand certain colloids may exercise a protective influence on other colloidal solutions. Thus, as Faraday first showed, colloidal gold is much more stable in the presence of a little gelatin. The colloids of serum can dissolve a considerable amount of purified globulin. Although the latter alone shows a drift in the electric field, the resulting solution is quite unaffected by the passage of a current through it. In these cases the protective colloids carry no charge, but the same protective effect may be observed if a large excess of *e.g.* an electro-positive colloid be added to an electro-negative colloid. This interaction between different colloids probably plays an important part in many physiological phenomena. We have reason to believe that the reactions between toxin and antitoxin and between enzyme and substrate, which we shall study later, are of this character, and that the compounds formed belong to the class of adsorption combinations.

In these combinations, although the chemical nature of the colloidal molecules is concerned, there is an absence of definite equilibrium points such as we are accustomed to in most chemical reactions. The inertia of the system and the large size of the molecules determine the occurrence of false equilibria and of delayed reaction, so that the condition and behaviour of a colloidal system at any moment are determined, not only by the quantitative relations of its components but also by the past history of the system.

THE COAGULATION OF COLLOIDS

Most colloidal solutions are unstable, and the relations between the suspended particle or molecule and the surrounding fluid may be upset by slight changes of reaction or the presence of minute traces of salts. As a result the hydrosol is destroyed, the suspended particles aggregating to form a precipitate or gel, the result varying according to the nature of the colloid and its concentration. Thus gelatin changes from the condition of hydrosol to hydrogel with fall of temperature. On the other hand, by adding calcium chloride to an alkaline solution of casein, we obtain a mixture which sets to a jelly on warming but becomes fluid again on cooling. Other agencies may lead to the production of changes which are irreversible.

If a solution of egg-albumin or globulin be heated in neutral or slightly acid medium and in the presence of neutral salt, the whole of it is thrown down in an insoluble form. This coagulated protein is insoluble in dilute acids or alkalies and is said to be *denatured.* With concentrated solutions of protein, heat coagulation results in the formation of a gel. In dilute solutions the result is the production of a flocculent precipitate.

Another method is the so-called mechanical coagulation. If a solution of globulin or albumin be violently shaken, a shreddy precipitate makes its appearance in the solution, and by prolonged shaking it is possible to throw down 80 or 90 per cent. of the dissolved protein in the coagulated form. Ramsden showed that this mechanical coagulation is a surface phenomenon. In the case of proteins the molecules at the surface are so closely packed together that they form an actual solid pellicle. When the solution is violently shaken, new surfaces are constantly being formed, and as the older surfaces are withdrawn into the fluid, the solid pellicle on them is rolled up into a fine shred of coagulated protein.

For Reference

ABRAMSON, MOYER and GORIN (1942). Electrophoresis of Proteins. (New York. Reinhold Publ. Co.)
BAYLISS, W. M. (1924). Principles of General Physiology. (Longmans, Green.)

Bayliss, W. M. (1923). The Colloidal State. (Oxford Medical Pub.)
Cohn, E. J. (1925). The Physical Chemistry of the Proteins. *Physiol. Review*, **5**, 349.
Hatschek (1925). An Introduction to the Physics and Chemistry of Colloids. (Churchill.)
Loeb, J. (1924). Proteins and the Theory of Colloidal Behaviour. (M'Graw Hill, N.Y.)
Luetscher (1947). Biological and Medical Applications of Electrophoresis. *Physiol. Rev.*, **27**, 621.
Pauli, W. (1922). Colloid Chemistry of the Proteins. (Churchill.)
Svedberg (1938). The Ultracentrifuge and its Field of Research. *Industr. & Engin. Chem. Analytical Edition*, **10**, 113.
Williams (1946). An Introduction to Chromatography. (London. Blackie & Son.)

CHAPTER VII

THE PASSAGE OF WATER AND SOLUTES THROUGH MEMBRANES

The heterogeneous structure of the body involves the presence of surfaces of separation, which may be of any order of size, molecular, microscopic, or evident on simple dissection. In the single cell the membrane bounding the cell and separating it from its environment is as a rule invisible under the microscope. Similar membranes must be assumed to be bounding the cell nucleus. In other cases the membrane is of a definite thickness: it may be apparently structureless and homogeneous such as the elastic lamina in the back part of the cornea of the eye, or it may be made up of cells of various shapes. The bounding membrane of the blood capillaries or of the lung alveoli is composed of very thin flat cells; or the membrane may consist of cubical or columnar cells, as in those forming the lining membrane of the small intestine or the essential part of secreting glands. In such a cellular membrane any transference across the membrane may be regulated by the cells themselves, taking up a substance on one side and turning it out on the other, in many cases after having undergone modification. It is impossible, however, to investigate the part played by these cells in the transference of substances across a membrane unless we know what factors determine the passage of water and dissolved substances across membranes which do not take an active part in the process of transference. Such a membrane we can imagine of any thickness, from one molecule upwards since the thickness will affect only the rate of passage of substances across it. Surface films, one molecule in thickness, are readily formed on liquids, *e.g.* when fatty acids are placed on water. In such a case the molecules are arranged vertically, with the carboxyl groups next to the water.* Similar *mono-molecular films* are also in all probability formed at the surface of colloid solutions and, as a general case, at surfaces where adsorption occurs.

The following factors may be concerned in causing the movement of water or dissolved substances across a membrane :—

(1) **Hydrostatic Pressure.** The significance of this is familiar in the ordinary process of filtration. A filter pump merely increases the difference of pressure between the two sides of the filtering membrane. The denser the membrane the greater will be the difference of pressure required to effect a passage at a certain rate.

(2) **Diffusion and Osmosis.** This has already been discussed in Chapter V.

The physico-chemical conditions for passage of water and salts across membranes have been studied in detail by Schreinemakers † and found to be complex, since when two solutes are present together in the solution on either side of a given membrane, the resultant transfer of water or of either solute depends on many factors. If the membrane is truly semi-permeable in the classical sense, *i.e.* is only permeable for water, equilibrium between two-solute mixtures will be reached when the two solutions have the same osmotic pressures, but their composition will probably not be the same. If the membrane is permeable to both solutes as well as to water, then when equilibrium is reached the composition on both sides will be identical, though the changes undergone in reaching such equality (or the "osmotic path") will be complicated and not predictable on general theoretical grounds, though modified by the properties of the membrane and the original composition and amounts of the solutions. A substance may at one phase of the diffusion pass in one direction and afterwards in the opposite one. In general terms, when n substances can pass a membrane, then there will be 2^n different types of diffusion, of which one is impossible, and

---

* ADAM. The Physics and Chemistry of Surfaces. 1930. (Clarendon Press, Oxford) LEATHES. *Lancet*, 1925, **103**, (1), 853.

† SCHREINEMAKERS. *J. Gen. Physiol.*, 1928, **11**, 710; 1929, **12**, 555; 1930, **13**, 335.

$2^n - 1$ possible. For instance, if there are six substances, there will be $2^6 - 1 = 63$ actually possible types of diffusion (*i.e.* when each substance may move either to right or to left that is counted as two types). Which of the possible types will actually occur depends on the properties of the membrane, and these cannot be predicted.

(3) **Electrical Difference of Potential.** This factor probably plays an important part in the transference of dissolved substances across membranes in the body. Just as differences of hydrostatic pressure determine movements of a fluid *en masse*, differences of electric pressure will serve as a driving force for the charged constituents of fluids. These may be actual particles, *e.g.* oil droplets or bacteria, the ultra-microscopic particles of colloids or the ions derived from the electrolytic dissociation of dissolved molecules. The passage depends on the fact that any charged particle will move towards the pole of opposite charge in an electric field. The electricity passing through the fluid is in fact carried by the particles which are thus set into movement. If a current is passed across a membrane immersed in an electrolyte, the final result will be a chemical difference on the two sides and therefore a permanent difference of potential which will tend, on breaking the current, to drive the ions in the opposite direction and to re-establish equilibrium on the two sides. The process of passage of a current across a membrane is said thus to give rise to *polarisation*, and after the passage of the current the membrane is polarised, *i.e.* presents a difference of potential between its two sides.

The presence of a membrane impermeable to the colloidal constituent of an ionised colloidal salt will affect the passage of a current through the membrane in one direction, so that the phenomenon of "irreciprocal conductivity" is presented. Thus if one side, A contains the potassium salt of hæmoglobin and the other side, B contains KCl, a current can pass from A to B, since it is carried from A to B by the potassium ions which can pass freely through the membrane, and from B to A by the electro-negative ions, namely, chlorine. An application of this method called *electro-dialysis*, is in fact one of the best ways of freeing protein solutions from adsorbed basic ions. If, however, B be made the positive pole, no current can pass because the electro-negative ion, namely hæmoglobin, cannot pass through the membrane. If we take an arrangement with three compartments A, B and C, of which only B contains a protein salt, the passage of a current will be hindered in both directions, since, whichever the direction of the current, the protein molecule will be required to carry the negative electricity either into A or into C. This resistance to the passage of a current will be still more marked if the membrane is impermeable also to certain inorganic ions. It is on this account that whole blood presents a greater resistance to the passage of a current than does the serum, each of the red blood corpuscles serving as a non-conductor and therefore diminishing the total area through which the current can pass.

In nearly every case the hindrance to migration of one or more ions by the membrane will prevent the passage of other associated ions to which by themselves the membrane would be freely permeable. These ions will travel an infinitesimal distance into the membrane, till they are held back by the electrostatic attraction of the oppositely charged ions which cannot pass. This electrostatic tension will give rise to a charge on the surface of the cell, and in nearly every case free swimming cells such as red blood corpuscles are found to be charged differently from the solution in which they are suspended. In most cases this charge is negative, so that, if a current be passed through blood, the red blood corpuscles will tend to travel towards the positive pole.

A charge on a membrane may also cause water to pass with an electric current through the membrane. This may be explained as follows: A charged particle, if free to move, will travel to the pole with the opposite charge; thus clay is negatively charged and migrates to the anode. Suppose now the membrane is made of clay and that electrodes are immersed in the fluid, the anode in A and the cathode in B. Now since the clay partition is fixed the water moves instead, viz. to the negative pole, so that the volume in B is increased. This phenomenon is called *electro-endosmosis*. With a gelatine membrane, the charge depends on the reaction of the solution. If this is alkaline, the gelatin is, like clay, negatively charged, but if acid it is positively charged and the direction of the flow is reversed.

(4) **Permeability of the Membrane.** In its relation to permeability, an animal membrane may be either passive or active.

(*a*) *Passive Permeability.* For this it makes no difference whether the membrane is dead or alive. The nature of the membrane cannot add to the force concerned in the transference of water or dissolved substances across it: it can only serve as a limiting condition for the action of the forces just enumerated. How these forces will act in regard to different kinds of membrane will depend on the permeability of the membrane, and this again on its physical and chemical structure. The membrane may be absolutely impermeable, as in the case of those surrounding certain eggs, and in plants as a means of protection for seeds. In most cases, however, the permeability is relative. Sometimes there is permeability to all crystalloids, as in a dialyser (*v.* Chapter V).

Some animal cells have a permeability of this dialyser-membrane type. Examples are the endothelial cells of capillaries in most parts of the body: the membranes formed by continuous sheets of such cells even allow a slight passage of the proteins of smaller size, especially when the endothelium is damaged. The clearest example of ultra-filtration is seen in the glomerular epithelium of the kidney. The capillaries of the central nervous system lie at the other extreme, and show a selective permeability towards many dissolved substances: water- and lipoid-soluble substances pass easily, anions, cations, glucose and urea only slowly.

A still greater restriction of permeability is obtained with "semi-permeable" membranes. No completely semi-permeable membrane is known, *i.e.* one which allows the passage of water but is impermeable to all dissolved substances. Such membranes are of much interest since, as already mentioned, the surfaces of all cells show in greater or less degree the phenomenon of semi-permeability. Thus, red blood corpuscles are only slowly permeable to most kations, such as Na and Ca, but allow the passage of K, urea, CO_2, chlorine, OH ions, as well as NH_4 and glucose (slowly).* To anions the order of permeability is $NO_3 > Cl >$ acetate > oxalate > sulphate > citrate, *i.e.* in order of increasing hydration.

Finally, we may have a form of permeability which depends on the solubility of the solute in the membrane itself, or on the formation of reversible compounds between the solute and the membrane. If a rubber balloon containing air be placed in a vessel filled with carbon dioxide, the balloon rapidly expands and may even burst. The carbon dioxide molecules are dissolved in the rubber and diffuse from the other side into the air inside the bag. The molecules of nitrogen and oxygen cannot pass in the reverse direction, so that the balloon expands. It has been supposed that phenomena of a similar nature may play a part in determining the peculiarly specific permeability of cell membranes. Cell surfaces are freely permeable to a number of organic substances, such as the alcohols which are soluble in lipides. It was suggested that the limiting membrane is lipoidal in nature, and hence allowed easy passage to lipoid-soluble substances, as already discussed on p. 16.

(*b*) *Active Permeability.* In many instances the plasma membrane is able to transport ions, and also non-ionised substances, from one side where a low concentration is present to the other side, where there is a high concentration. This transport, against the gradient of concentration must involve the utilisation of energy, provided by the metabolism of the cell contents, and is therefore dependent upon the life of the cell.

One way in which transport could be effected is by change of the material transported into some other substance, which is not diffusible, *e.g.* if a cell can convert glucose into glycogen, or phosphate ion into creatine phosphate or A.T.P. This would, however, although dependent on the metabolism of the cell, be a different case, and not an example of transport against a gradient.

It is different with most inorganic ions, for these are apparently all present in the free state within the cell as well as in its environmental fluid. Yet in almost every type of cell, plant or animal, we find that the ionic contents of the cell differs widely from that of the fluid outside it in which it has grown and lived. Thus the Cl ion in the frog muscle fibre is only 1·2 m Eq/l, while that in blood serum is 74·3 m Eq/l. The red blood cell is another example of different ionic distributions inside and outside, as shown in the table (p. 82).

In each case the concentration of K is higher, sometimes very much higher, and that of Na lower, inside the cell than outside. It has often been supposed that, because this difference is maintained, the cell membranes must be impermeable to one (Na) or both ions. But experiments with the radio-active isotopes ^{42}K and ^{24}Na do not support this view. During muscular work some K leaves the muscle fibres and is replaced by Na. Also, in absence of glucose, K leaves the red blood cells and returns again when glucose is added. We must

* Hsu. *J. Physiol.*, 1935, **84, 29P**.

Cell.	Concentration in mm/1. of Fluid.			
	Potassium.		Sodium.	
	Inside.	Outside.	Inside.	Outside.
Dog, red blood cell .	9	5·5	145	147
Man, red blood cell .	125	5·9	36	160
Frog, muscle . .	126	2·6	15·5	103·8

conclude, therefore, that these ionic transportations are effected as a result o the expenditure of energy by the cell.

The permeability of living membranes often shows peculiarities, which depend on thei being alive. For instance, Reid * found that fresh frog's skin allows water to pass only from without inwards, and Wertheimer,† found that in frog's skin sodium salts and acid dyes only pass from without inwards, while glucose and basic dyes only pass outwards.

The explanation of irreciprocal permeability probably varies from instance to instance, but is mostly unexplained. As regards water, it is thought to be related to the so-called anomalous osmosis which depends on the membrane being electrically charged. Thus, the frog skin has a negative charge outside and a positive one within : this charge is dependent on the metabolism of the skin, as a result of which H′ and HCO_3′ are kept at a high level on the inside. The resultant water movement is probably due to electro-osmosis.

The surface membrane of a cell differs, moreover, from artificial semi-permeable membranes in that its permeability varies according to the state of the cell itself. Thus, activity of a cell is always attended with an increased permeability of its surface membrane, and this change may be an important factor in the execution of its functions.

Not only does the function of the cell determine the properties of its surface, but the reverse holds true. Cell activities seem to be determined to a very large extent by alterations in the amounts of substances in the surrounding medium even though they cannot penetrate the limiting membrane of the cell. This is especially well marked in the case of contractile tissues, such as the heart, where the normal beat depends on the presence in the surrounding fluid of K, Na and Ca ions. None of these can penetrate the cell membrane : they can merely be adsorbed on the cell surface, and it is their presence there which in some way or other modifies all the functional activities of the muscle cell.

(5) **Membrane Equilibria.** If a colloid is in the form of a compound with either acid or base, which undergoes ionic dissociation, certain interesting results are possible.‡ For instance, Congo red is a sodium salt of a colloid acid. If this be placed in an osmometer which is immersed in water, a certain osmotic pressure is developed. Although sodium ions can pass readily through the membrane of the osmometer, they carry a positive charge, so that if they broke away from their combination with Congo red acid, the fluid in the osmometer would become negatively charged. The sodium ions are thus held in the immediate neighbourhood of the membrane by the electrostatic attraction between Congo red acid ions and the sodium ions. If we add sodium chloride to either the inner or outer fluid, the osmotic pressure is found to fall if time be allowed for equilibrium to be established between the two sides of the membrane. At this point the outer fluid, which is free from dye, contains a larger percentage of chloride than the inner fluid. This difference is permanent and is more marked the greater the concentration of the dye. In the following Table are given the concentrations of the two fluids with different percentages of salt, the numbers indicating the litres to which each gramme molecule of the salt is diluted (i.e., the *reciprocals* of the the concentrations).

Dye.	Chlorine.		
	Inside.		*Outside.*
30	52	...	30
30	465	...	73·6
30	5500	...	180
100	32·9	...	29·5

---

* Reid. *J. Physiol.*, 1890, **11**, 312.

† Wertheimer. *Pfluger's Arch.*, 1923, **199**, 383.

‡ Bayliss, W. M. *Proc. Roy. Soc.*, 1909, **81B**, 269 ; 1911, **84B**, 229.

Donnan * independently showed that these results must occur from thermodynamic considerations, so that their general statement is spoken of as the "Donnan equilibrium."

When a saline solution contained in a collodion sac is suspended in distilled water, the simple laws of diffusion suffice to explain the final equal concentrations of salt reached on both sides of this permeable membrane. If, however, a sodium salt of a non-diffusible anion (R′) be placed on one side of a dialysing membrane and water on the other side, simple osmosis does not suffice to explain the final state of equilibrium. The Na will tend to pass through the membrane, but this can only take place if an equivalent quantity of OH (from the water) diffuses in the same direction. Thus :

Initial State		*Final State*	
$Na^{\cdot}$	H_2O	$Na^{\cdot}$	$Na^{\cdot}$
R'		$H^{\cdot}$	OH'
		R'	

The water will become alkaline, while the solution in the left compartment will become less alkaline. If, on the other hand, a salt, say a chloride, of a non-diffusible kation, were placed on the left-hand side, then the water outside would become acid, and the solution on the left less acid.

Again if NaR be placed on one side of the membrane and NaCl on the other side, equilibrium will be attained only when the products of the concentrations of [Na] and [Cl] are the same on opposite sides of the membrane.

Initial State		*Final State*	
$Na^{\cdot}$	$Na^{\cdot}$	$Na^{\cdot}$ Cl'	$Na^{\cdot}$ Cl'
R'	Cl'	R'	

$$[Na]_1 \times [Cl]_1 = [Na]_2 \times [Cl]_2$$

This can only occur if $[Na^{\cdot}]_1 > [Na^{\cdot}]_2$ and $[Cl']_1 < [Cl']_2$.

This difference in the concentration of the diffusible ions on opposite sides of the membrane gives rise to a potential difference which is equal to

$$58 \log. \frac{[Cl']_1}{[Cl']_2} \text{ millivolts.}$$

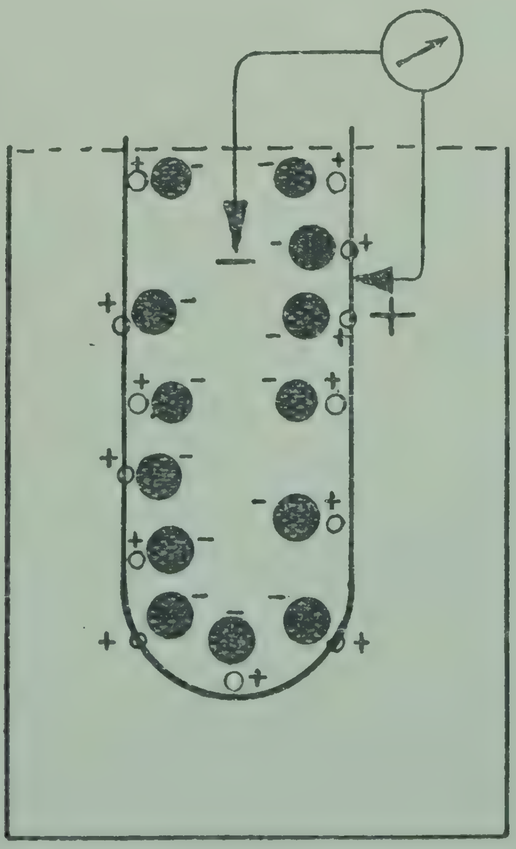


FIG. 26.

---

* DONNAN. *Ztschr. f. Elektrochem.*, 1911, **17**, 572.

Thus, suppose we have, as in Fig. 26, a collodion sac filled with an alkalin solution of a protein, *i.e.* with the protein in the form of a salt with a base, an immersed in water or saline solution.

The proteinate solution will be ionised into the basic ion B^+ and the larg protein acid ion Pr^-, and the latter cannot pass through the collodion therefore, although the collodion is permeable to the basic ion, this cannot pas through it either, because it is kept back by the attraction of the oppositel charged protein ion. But, owing to the tendency of the B^+ ion to penetrat the membrane, the two ions are arranged in an orderly fashion, the negativel charged ones innermost and the positively charged ones next to the membran or perhaps just within its substance. The membrane is thus *polarised* o charged, the outer surface positively, so that if we lead off from its outer an inner surfaces a potential difference can be demonstrated and shown to agre with theoretical calculations.

Since living cells are likewise filled with protein and other colloidal substance in combination, partially ionised, and the living tissues abound in membrane impermeable to colloids, Donnan's theory of "membrane equilibria" must have a wide application in physiology. But the effects are modified, and often obscured, as a result of other properties of the cell membranes, *e.g.* that of being permeable to some ions and not to others. Thus red blood cells are permeable to anions but not to cations ; their surface is negatively charged.

For Reference

BOLAM (1932). The Donnan Equilibria. (Bell & Co.)

DAVSON and DANIELLI (1943). The Permeability of Natural Membranes. (Cambridge Univ. Press.)

GELLHORN and RÉGNIER (1936). La perméabilité en physiologie et en pathologie génerale. (Paris.)

HITCHCOCK (1924). Proteins and the Donnan Equilibrium. *Physiol. Rev.*, **4**, 505.

HÖBER (1936). Membrane Permeability to Solutes. *Physiol. Rev.*, **16**, 52.

JACOBS (1939). Permeability. *Ann. Rev. Physiol.*, I., 1.

KROGH (1946). The Active and Passive Exchanges of Inorganic Ions through the Surfaces of Living Cells and through Living Membranes generally. *Proc. R. S.*, **133B**, 140.

CHAPTER VIII

BIOELECTRIC POTENTIALS

It has long been known that all living tissues, animal and plant, are capable of showing developments of electric potential. In the exceptional cases of the electric fishes potential differences of several hundred volts, capable of stunning a horse, can be generated, but in most cases the potential is measurable in millivolts, and rarely exceeds 100 mV. Nevertheless these potentials are easily measured and recorded and have been the subject of much attention for many decades. They are of especial interest in connection with the active state of tissues, *e.g.* contraction of muscle, impulse transmission in nerve, secretion of gland cells, and are believed to be fundamental phenomena of the excitatory state to which those activities are due.

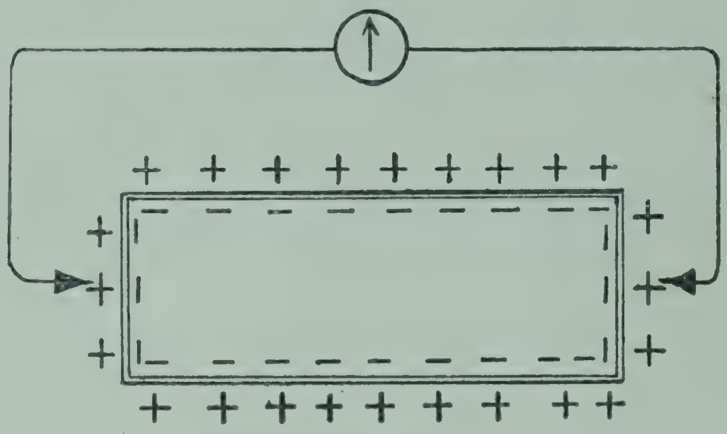


Fig. 27.

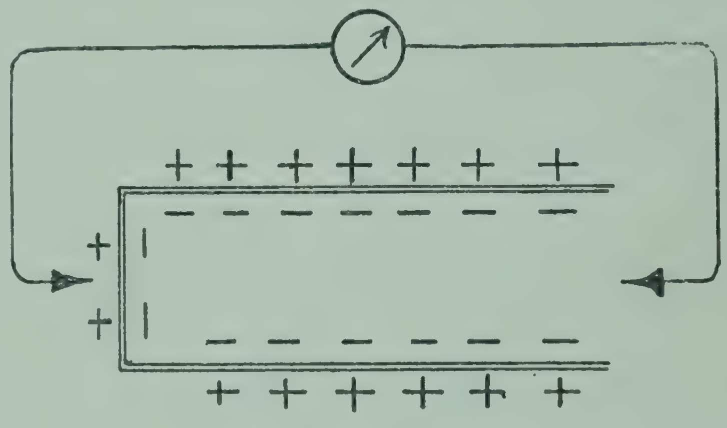


Fig. 28.

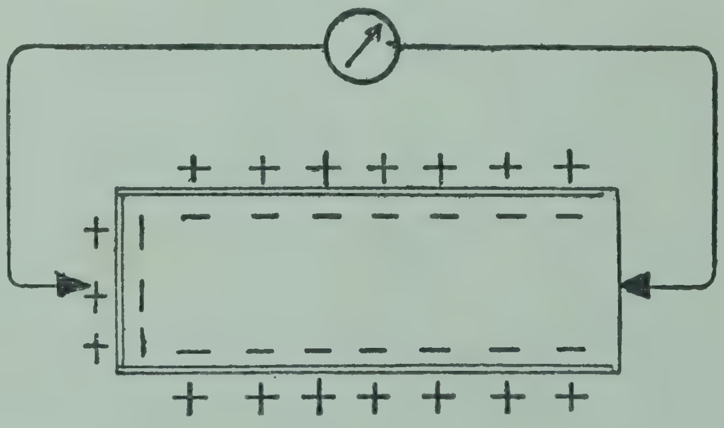


Fig. 29.

Figs. 27, 28, 29. To show the origin of bioelectric potentials. The double line round the cell indicates a polarised membrane, the single line in Fig. 29 a depolarised one.

In general, any two points at the surface of a resting cell are at the same potential, and the resting cells in a tissue all have the same surface potential. When a cell, or part of it, is excited, however, the surface of that part which is excited becomes negative to the resting surface, so that if contact between active and resting surfaces is made through a galvanometer, a current will flow from the resting to the active tissue, as long as the activity lasts. Further, if the surface of a resting cell is broken at any point, a lead-off from the damaged area, which is equivalent to a lead-off from the cell contents, will show this also to be negative to the resting and uninjured outer cell surface.

These potential differences must be due to unequal distribution of ions, and much convergent evidence leads us to regard the plasma membrane as the principal theatre of this ionic orientation : the bioelectric potentials are hence often spoken of as *membrane potentials*, and are considered as being due to *polarisation* of the plasma membrane. By this we mean that attached to, or incorporated in, the inner surface of the membrane is an excess of negatively charged anions, while on the outer surface are either very few anions, or else an excess of cations. Hence the membrane somewhat resembles a charged condenser, with the electrons accumulated on the inner surface, but itself of high insulating power so that the charge is maintained (Fig. 27). Damage to the membrane enables the two surfaces to be connected through a galvanometer, so that current flows (Fig. 28). But it is a continuous, not a brief current, that the charge on the membrane must be maintained in some way, and t analogy with a condenser can only be a partial one.

When a part of the cell becomes active, current also flows in a simil direction, *i.e.* from the resting to the active surface (Fig. 29), and it can be show that at the same time the resistance of the membrane at the active part

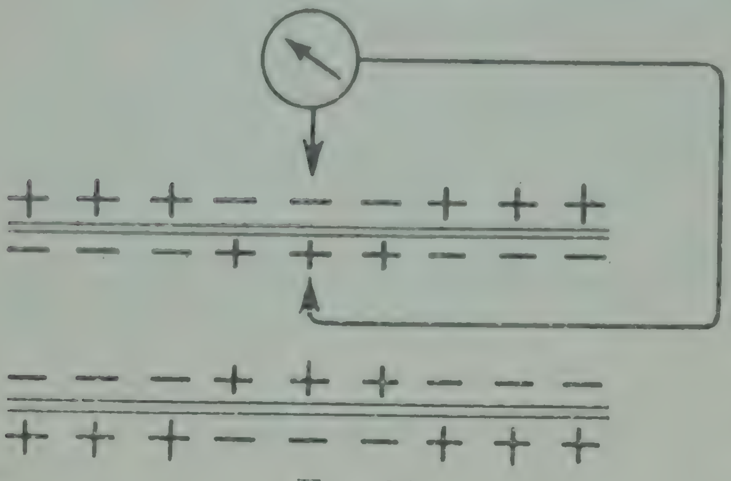


Fig. 30

suddenly lowered to about $\frac{1}{40}$ of i resting value, allowing the current flow. Activity has thus been regarde as due to local depolarisation of th plasma membrane. But it now appear that it is something more than mer local *de*polarisation, and that, in fac the polarity of the membrane, is actuall *reversed*, *i.e.* there is a local excess o negatively charged ions on the *outsid* instead of on the inside of the membrane; this has been shown, as i Fig. 30, by leading off from both sides of the membrane with micro-electrodes, avoiding all short circuiting.

The cause of the polarisation of the resting plasma membrane is not exactly known, but is probably related to a specific ionic permeability, due to the molecular structure of the membrane. The potassium ion is small, and can enter the membrane; the sodium ion is hydrated and thus is too large to penetrate. We have seen (p. 17 and Fig. 9) that orientated lipoid molecules are probably the basis of the plasma membrane: the orientation is such that the polar groups of the lipoid molecules are at or near the surfaces and are so turned that their negatively charged ends are at the surfaces and the positively charged poles within the membrane as in Fig. 31. These negative charges attract the cations in the aqueous phase, of which potassium only can penetrate, however: the potassium does not pass right through the membrane, but merely enters the surface layer next to the negative charge on the dipole, and when it does so the anion is left behind, but remaining at or near the surface owing to the attraction exercised by the adsorbed cations. The surface of the membrane thus acquires a negative charge, the value of which depends on the concentration

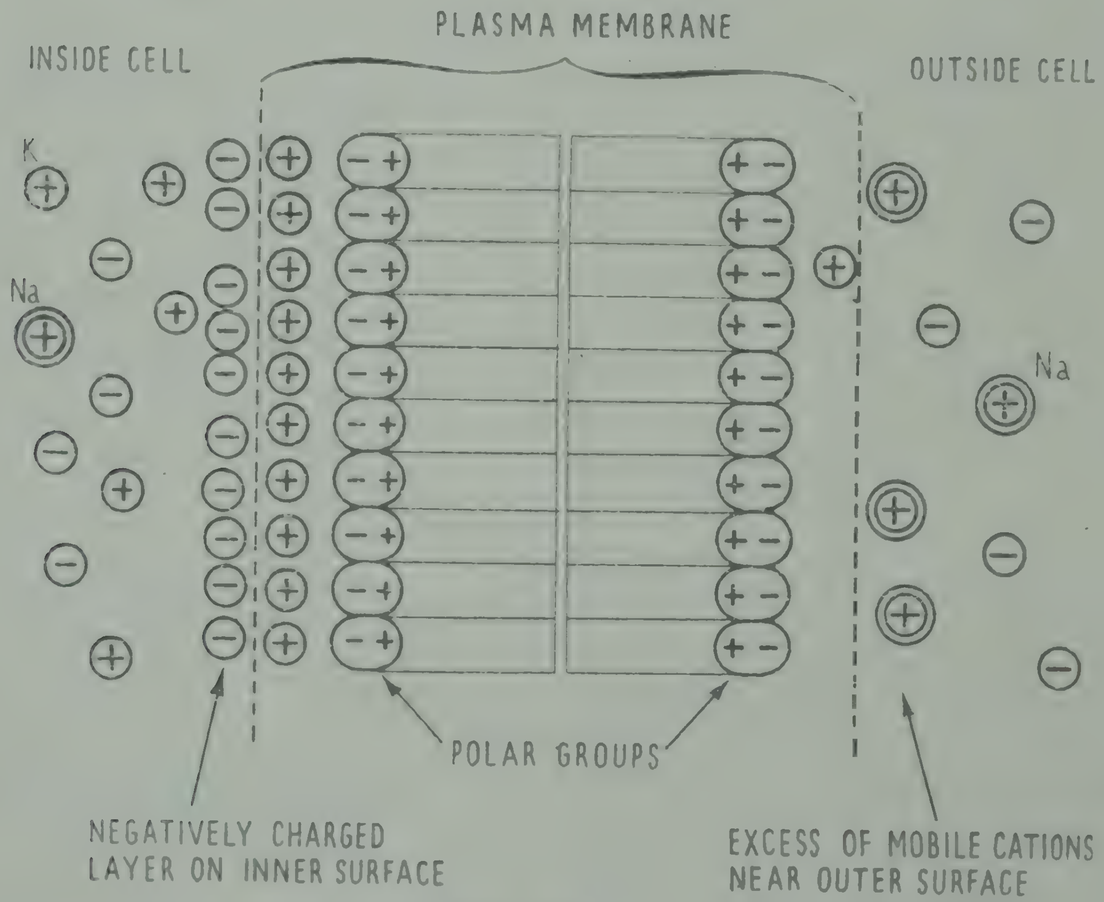

PLASMA MEMBRANE
INSIDE CELL
OUTSIDE CELL
K
Na
Na
POLAR GROUPS
NEGATIVELY CHARGED LAYER ON INNER SURFACE
EXCESS OF MOBILE CATIONS NEAR OUTER SURFACE

Fig. 31

of potassium in the solutions on the two sides. In the excitable tissues, there is a high concentration of K inside the cell and a low one outside, where Na replaces it (for squid fibres the potassium ratio is 279 m Eq/l. in axoplasm to 10 m Eq/l. in sea-water (and tissue fluid) *) ; hence the inside of the membrane is negative to the outside, as we have seen.

The maintenance of the high concentration of K ions and the low level of Na ions inside the cell is due to the expenditure of energy by the cell, and ceases when the cell dies, or when it is deprived of oxygen. It is therefore the result of metabolic changes taking place in the cell.

During the state of excitation the permeability of the plasma membrane is altered. According to the classical theory it now becomes freely permeable to both anions and cations, as a result of a molecular breakdown, and hence the active region of the membrane loses its charge ; it is then *relatively* negative to the resting surface. But, as stated above, more recent work has shown that, at all events in some instances, the charge on the membrane is not merely lost but reversed, so the depolarisation theory now appears inadequate. What actually happens is a matter for further conjecture. An attractive suggestion is that the inner one of the two orientated lipoid layers becomes disorientated, leaving the outer layer unopposed and therefore with a considerable negative charge at the surface.

---

* Webb and Young. *J. Physiol.*, 1940, **98**, 299.

CHAPTER IX

ENZYMES AND ENZYME ACTION

Most of the phenomena of life are accompanied and conditioned by chemica changes, both of synthesis and of disintegration, of the most varied character Synthesis by ordinary laboratory methods falls far short of the process as i actually occurs in the living cell, both in completeness of the reaction and in th ease with which it can be effected. Many of the chemical transformation effected by living cells we have so far been quite unable to imitate. Th distinguishing characteristics of the net chemical changes wrought by th living cell are :

(1) The rapidity with which they are effected at relatively low temperatures.

(2) The specific direction of the process, which is therefore almost complete, with a surprising absence of side reactions. This characteristic may, however, be regarded as a consequence of the first, since an increase in the velocity of any given reaction will determine a preponderance of this reaction over all others.

(3) The usage, or production, of optically active substances when possible.

All chemical changes in living beings take place in a watery medium. But water plays a direct part in many of the reactions which take place. Apparently the molecule may act as a whole, or the H and OH ions may be separately involved. When the molecule acts as a whole, *e.g.* in hydrolysis, the water is added to the molecule on which it acts, with the production of a hydrated product or with a splitting of the original molecule into two parts as a result. The rapidity with which these changes occur in the body is very striking. A solution of boiled starch in water may be left for many months without undergoing change, but if we take the solution into the mouth, keep it there for twenty seconds and then eject it, the fluid is found to contain no starch but to contain maltose. In this case there has been a taking up of water molecules by the starch, so that the interaction of water appears in the equation—

$$2\ (C_6H_{10}O_5)_n + nH_2O = nC_{12}H_{22}O_{11}$$

This change, which would never occur spontaneously, is due to the action of a catalytic agent present in the saliva. Such agents are given the general name of *enzymes*, and they play a dominant part in the carrying out of the varied chemical changes which occur in all living things.

ENZYMES

After 1857, when Pasteur showed that the process of alcoholic fermentation, by which glucose is converted into alcohol and carbon dioxide, was apparently inseparable from the presence and growth of living yeast cells, it became customary to suppose that most of the chemical changes which took place in all living things were in some similar way dependent upon the activity and structural integrity of living cells. But in 1897 the Buchner brothers showed that a cell-free juice obtained from yeast could effect an identical fermentation of glucose. We now know that this yeast juice contains a mixture of enzymes by which the chemical reactions known as fermentation are catalytically promoted ; the present-day view is that nearly all the chemical reactions of life are similarly promoted by enzymes, which, when isolated, can perform the same catalyses *in vitro*. Minimal quantities of enzyme can induce chemical changes involving almost unlimited quantities of other substances, the only result of increasing the quantity of enzyme being to quicken the rate of the change. They are, in fact, *catalysers* or *catalysts*. A catalyser is a substance which will increase the velocity of a reaction without adding in any way to the energy changes involved n the reaction, or taking any part in the formation of the end products. The eactions catalysed by each enzyme are very restricted in character, *i.e.* each nzyme is specific. The reactions are usually reversible.

Enzymes of various kinds occur in the digestive juices. Others may often e extracted from the tissues with water, glycerol, salt solutions, or dilute lcohol. Others are insoluble, either because they are bound chemically to the ell cytoplasm, or because they are adsorbed to a constituent of the cytoplasm. he difficulty in their isolation is increased by the fact that all of them are olloidal, and present the tendency common to all colloids of being adsorbed o other colloidal matter as well as to precipitates. A common method of btaining a concentrated preparation of an enzyme is to add to its solution strongly adsorbing colloidal precipitate, such as alumina or kaolin. Kaolin, eing negatively charged, adsorbs those enzymes which have basic characters, while alumina, a positively charged colloid, adsorbs acid substances. The nzyme is carried down on the adsorbent and may be obtained in solution by "elution" with water or dilute acid or alkali. By fractional adsorption with different adsorbents, enzymes may be greatly concentrated and also separated from one another. In this way many enzymes formerly considered s entities have been shown to consist of mixtures of two or more enzymes, and some have been obtained in a pure crystalline state.

All the known enzymes are proteins, or conjugated proteins.

The substance which is chemically altered as a result of enzyme action is called the *substrate.*

It is important to note that nearly all the changes brought by enzymes can be effected by ordinary chemical means. Thus the disaccharides can be made to take up a molecule of water and undergo conversion into monosaccharides. A solution of maltose at ordinary temperatures undergoes practically no change. If the solution be warmed, a slow process of hydration takes place which is quickened by rise of temperature, so that if the solution be heated under pressure to, say, 110° C., hydrolysis occurs with considerable rapidity. If, however, a little maltase be added to the solution, the change of maltose into glucose takes place rapidly at a temperature of 30° C. The action of the enzyme is to quicken a process of hydrolysis which without its presence would take an infinity of time for its accomplishment.

A very large number of enzymes are now recognised, and new additions are constantly being made to the list as our knowledge of cellular metabolism expands.

The enzymes are often remarkably specific in the substrates on which they can act. Whereas all the disaccharides are converted by acids into the corresponding monosaccharides, the enzyme invertase acts only on cane sugar and has no action on maltose or lactose, each of which requires a specific enzyme (maltase, lactase) to effect its hydrolysis. But we can also find examples of a restricted action even among inorganic catalysers.

In the chemical transformations of cellular metabolism each reaction is usually the result of the action of a specific enzyme which catalyses that reaction, and that reaction alone. This high specificity represents an organisation of the functions of the living cells at the molecular level. It is clear that the chemical changes which take place in the course of metabolism must be the result of the action, in an ordered sequence, of large numbers of enzymes.

Absolute specificity is shown as between enzymes of different groups : *e.g.* lipases will not act on proteins. In other directions, within their groups, the specificity is variable, and is probably related to particular chemical groups in the substrate : thus lipases will also hydrolyse many other esters besides fats. It has been found that, whereas maltase splits up all the α-glucosides, it has no power on the β-glucosides : that is to say, maltase will fit into a molecule of a certain configuration, but is powerless to affect a molecule which differs from the first only in its stereochemical structure. On the other hand emulsin which breaks up β-glucosides, has no influence on α-glucosides. This specific affinity of the enzymes for optically active groups of bodies suggests that the operative groups of the enzyme itself may be optically active.

THE NATURE OF ENZYMES. By the use of selective adsorption, enzymes may separated from one another and from many accompanying impurities, and obtained i state of great concentration. Crystalline preparations of several enzymes, *e.g.* urease pepsin[†], chymotrypsin, trypsin[‡], amylase, catalase, etc., have been obtained.

All enzymes are colloidal; some at all events consist of two parts closely associat viz., a colloidal "carrier" which is non-specific, and one or more specific active groups, which union with the substrate is effected. This union takes place with definite groups the substrate molecule, and the specific action of the enzyme is determined by the natu and position of the receptor group to which the specific group of the side chain becom attached. Thus, in the enzymes of the peptidase class the imino group is one point attachment. The purest enzymes, crystalline pepsin[§], trypsin[‖], urease[¶] and amylase are proteins, and probably all enzymes are of protein nature, or contain a protein as essential constituent. Some (*e.g.* urease, dehydrogenase, lipase, phosphatase) depend f their activity on possession of an SH group; such enzymes are inactivated by oxidisi agents which convert the SH groups to S–S groups, or by iodoacetates, which combine wi the SH group; others, *e.g.* pepsin depend on the intactness of the hydroxyl groups tyrosine residues. Many enzymes can exist in an inactive form which is convertible, slight molecular rearrangement, into the active form, the conversion in some cases, *e.g.* the activation of trypsin, being autocatalytic, *i.e.* it is started by a minute trace of t active form, and proceeds with rapidly augmenting velocity as the active form increas in amount.[††]

Enzymes as Catalysers. The subject of thermodynamics teaches tha no reaction can be brought about through the agency of a catalyst unless it accomplishment is attended by a diminution of free energy (*v.* p. 67) of th whole system under consideration. The matter may perhaps be illustrated by mechanical analogy. Suppose a 100 g. weight to be placed at one end of a shee of glass, which is sloped so that the weight is 10 cm. from the table surface an the other end is at the table surface level. Friction, let us suppose, prevent the weight sliding down the inclined surface, although the weight has potentia energy capable of doing 1,000 g.-cm. of work, *i.e.* if it did slide down ΔF woul be negative. If now a little oil is placed beneath the weight, it slides down t table level, and in so doing could perform work and/or generate heat which together would be equivalent to 1,000 g.-cm. This energy liberation would b the measure of the diminution in the free energy of the change, and the oil which merely removes the frictional resistance to it, corresponds to the catalyst. (If the weight then rolled on to the floor, there would be a further liberation of free energy, and if it next rolled down the staircase there would b a still further diminution in free energy. But we are only concerned here with the move down to table level. In other words we are only concerned with *changes* in free energy in particular reactions.)

Now it is self-evident that if we want the weight to go up the inclined plan we must push or lift it up, *i.e.* we must add energy to the system in some wa or the reaction will never occur, oil or no oil (*i.e.* ΔF is positive, *v.* p. 67).

Enzyme and other catalysed reactions are similar; they will not occur unles accompanied by a diminution of free energy. The catalyst, like the oil, merely hastens or makes possible a reaction which is possible on energy grounds, an at the end is unchanged itself. Whereas the influence of most catalysers o the velocity of a reaction increases rapidly with rise of temperature, in the cas of enzymes this increase occurs only up to a certain point (the temperatur coefficient varying, up to this point, from 1 to 2). This point is spoken of as th *optimum temperature* of the enzyme action. If the mixture be heated above thi point, the action of the enzyme rapidly slows off and then ceases, owing t denaturation of the enzyme.

All enzymes are very sensitive to changes of hydrogen ion concentration and there is an optimum *p*H at which each enzyme is most active, *e.g.* for pepsin

[*] SUMNER. *J. Biol. Chem.*, 1926, **69**, 435; **70**, 97.

[†] NORTHROP. *J. Gen. Physiol.*, 1931, **14**, 713.

[‡] NORTHROP and KUNITZ. *J. Gen. Physiol.*, 1932, **16**, 267.

[§] NORTHROP. *J. Gen. Physiol.*, 1930, **13**, 739.

[‖] NORTHROP and KUNITZ. *J. Gen. Physiol.*, 1932, **16**, 267.

[¶] SUMNER, KIRK and HOWELL. *J. Biol. Chem.*, 1932, **98**, 543.

[**] CALDWELL and SHERMAN. *Science*, 1931, **74**, 37.

[††] NORTHROP. *Science*, 1937, **86**, 479.

*p*H 1·5, for trypsin *p*H 8·3, etc. The optimal *p*H, however, depends very much on other conditions, *e.g.* in the presence of impurities it may be considerably altered.

Many enzymes are easily inhibited or destroyed by certain substances, *e.g.*, by the salts of Hg, Cu, Pb, by iodine, some by HCN, etc., in very small concentrations.

KINETICS. Since the action of enzymes consists in the quickening up of processes which would otherwise occur at an infinitely slow rate, many investigations have dealt with the velocity of the reaction induced by the action of an enzyme. It has been found that the velocity of an enzyme action is accelerated by rise of temperature (up to the optimum) by an optimal *p*H, and by increase in concentration of substrate or of enzyme. Conversely it is inhibited by accumulation of the end-products of the reaction, and hence the action is probably a special case of the laws of mass-action ; it may also be slowed down by inactivation of the enzyme.

Enzymes, which are all colloidal, cannot be dealt with in the same way as the simpler catalysts. In many cases the substrate, *e.g.* starch or protein, is also colloidal, and the combination therefore falls into the class of combinations between colloids. In this we have an interaction between two substances in which the adsorption by the surfaces of the molecules of one or both substances plays an important part.

A variety of methods are available for studying the kinetics of enzyme action. The general principle of these is that the amount of substrate broken up by the enzyme is determined by some appropriate chemical or physical method which measures either the concentration of the substrate or of one of the products at a given time after the enzyme and substrate were brought together. Factors which may be varied are concentration of enzyme and of substrate, time of action, temperature, H-ion concentration, and presence of various other substances.

By such methods it has been found in the simplest case, when the substrate is in true solution, that, when the quantity of enzyme employed is very small in comparison with the *substrate* (the substance acted upon), the *amount of change in a given time is proportional to the amount of enzyme* present, and is (within limits) independent of the concentration of the substrate. This is well shown by the two following Tables representing the action of lactase upon lactose (E. F. Armstrong *) :

PROPORTIONS HYDROLYSED IN 100 C.C. OF A 5 PER CENT. SOLUTION OF LACTOSE. ENZYME VARIED.

Solutions containing—	1·5 hours	20 hours	45 hours
1 c.c. lactase . . .	0·15	2·2	3·9
10 c.c. ,, . . .	1·6	23·3	38·6
20 c.c. ,, . . .	3·2	45·8	—

LACTOSE HYDROLYSED. ENZYME CONSTANT, SUBSTRATE VARIED.

Solutions containing—	24 hours		46 hours	
	Proportion	Weight	Proportion	Weight
10 per cent. lactose .	14·2	1·42	22·2	2·22
20 ,, ,, .	7·0	1·40	10·9	2·18
30 ,, ,, .	4·8	1·44	7·7	2·21

---

* E. F. ARMSTRONG. *Proc. Roy. Soc.*, 1904, **73**, 500.

Moreover if we take only the earlier stages of the enzyme action, it is found that, wi small proportions of enzyme, equal amounts of substrate are changed in successive interva of time until about 10 per cent. has been hydrolysed. This is shown in the followi Table :

2 PER CENT. LACTOSE WITH LACTASE

Time	Amount hydrolysed
⅓ hour	3·2
⅔ ,,	6·4
1 ,,	9·6
2 hours	16·4
3 ,,	20·8

These results can best be interpreted by assuming that the first stage in the reaction is combination of enzyme with substrate. Bayliss believed the combination to be an adsorp tion, whereas Michaelis thought it to be a reversible combination of the ordinary chemic equilibrium type. In either case, it is the concentration of this combination which represent the active mass of the molecules, *i.e.* the molecules of substrate which are undergoin change. This compound, as soon as it is formed, takes up water and breaks down, settin free the hydrolysed substrate and the enzyme, which is at once ready to combine with further portion of the substrate. In such a case the velocity of reaction must be directl proportional to the amount of enzyme, and the same absolute quantity of substance wi continue to be changed in succeeding units of time. Supposing, for instance, we had a the bottom of a hill a load of bricks which had to be transferred to the top, and some men t effect the transference. The rate of transference would be directly proportional t the number of men employed ; we could double the rate by doubling the men. Moreove the number of bricks carried in each unit of time would be the same. Five men would carry as many bricks in the second ten minutes as they would in the first, and so on. On the othe hand, the velocity with which the transference was effected would be independent of the number—that is, the *concentration*—of the bricks at the bottom of the hill. The *active mass* of bricks could be regarded as that number carried at any moment by the transferring factor, namely the men. The equation of change would be $\phi = KC$, where C is the concentration of the enzyme. This concentration is always being renewed and kept constant by the breaking down of the intermediate product, so that the rate of change would be continuous throughout the experiment.

On the other hand when the amount of enzyme is relatively large, the rate of change, though at first very rapid, tends continuously to diminish. This is shown by the following Table representing the rates of change (K), during succeeding intervals of ten minutes, in a caseinogen solution to which a strong solution of trypsin had been added (Bayliss).

VELOCITY OF TRYPSIN REACTION

6 c.c. 8 per cent. Caseinogen + 2 c.c. $\frac{N}{10}$ NH_4OH + 2 c.c. 2 per cent. Trypsin at 39° C

1st 10 minutes	K = 0·0079
2nd ,,	0·0046
3rd ,,	0·0032
4th ,,	0·0022
5th ,,	0·0016
6th ,,	0·0009
&c.	&c.

In this retardation of the later stages of enzyme action, the end products are concerned in some way or other, and the retardation can be augmented by adding to the digesting mixture the boiled end products of a previous digestion. The retarding effect of the end products resembles in many ways that observed in reversible reactions and is most simply regarded as being due to the fact that the enzyme reversibly combines not only with the substrate, but also with the end products, tending to reverse the reaction, or to prevent enzyme from combining with substrate, or both.

In general, however, reversal of enzyme actions, though theoretically sound, is not easily attained, so that probably a more important factor in slowing the enzyme action in its final stages is the preliminary combination of the enzyme itself with the end products and the consequent removal of the enzyme from the sphere of action. Several facts speak for such a mode of explanation. Thus the action of lactase on milk sugar is not retarded by both its end products, glucose and galactose, but only by galactose. In the same way the action of invertase on cane sugar is retarded by the end product fructose but not by glucose. These are examples of "*competitive inhibition*," one sugar seizing the enzyme and excluding the other.

So far, therefore, a study of the velocity of enzyme actions would lead us to suspect that the enzyme in the first place combines with the substrate, and that this combination is a necessary step in the alteration of the substrate. In the second place, the enzyme is taken up, to a certain extent, by some or all of the end products, and this combination acts in opposition to the first combination, tending to remove the enzyme from the sphere of action, and therefore to retard or even reverse the whole reaction.

Activators and Co-enzymes. Some enzymes are quite inactive except in the presence of some other substance or substances, which are called *activators* or *co-enzymes*, the differences between the two being rather vague. Thus salivary amylase is inactive in absence of chloride or bromide ions, and such simple accessories are commonly called *activators*.

A different kind of activation is that known as "*unmasking*." For instance, pancreatic juice contains a precursor of the proteolytic enzyme trypsin, but it will not act on proteins until mixed with intestinal juice. This contains a substance called *enterokinase*, which appears to remove part of the large trypsinogen molecule and so expose the enzyme trypsin ready for action.

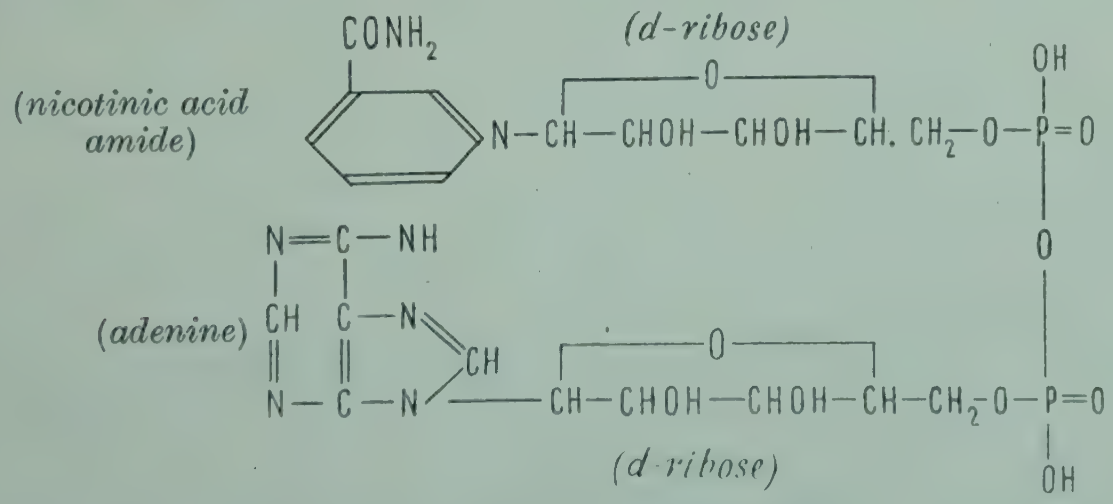
Yet another type of activation is seen in the operation of co-enzymes. For example, an extract of muscle can be made containing various enzymes which, acting together, can convert glycogen into lactic acid. But if the extract is dialysed it will not do this. Addition of the dialysate, or of boiled extract, restores the property : there is evidence that the essential substances removed by dialysis play a *rôle* in the reactions carried out by the enzymes, and such participants are called *co-enzymes*, the complete system consisting then of enzymes and co-enzymes. Sometimes more than one accessory substance is essential, and these together are then called a *co-enzyme system*. Enzymes may usually be freed from their co-enzymes, or from the various components of their co-enzyme systems, by dialysis. The co-enzymes are usually thermostable. There is evidence that the co-enzyme combines with the enzyme, either temporarily, much in the way that the substrate does, or more permanently in the form of the prosthetic group of a conjugated protein.

The following are important co-enzymes of known constitution :—

Co-enzyme		*Example of Enzyme using*
(1) Co-enzyme I	Phospho-pyridino-nucleotides.	Lactic dehydrogenase.
(2) Co-enzyme II		*iso*-citric dehydrogenase.
(3) Flavine nucleotides.		some Reductases.
(4) Thiamine pyrophosphate.		Carboxylase.
(5) Iron-porphyrin compound.		Cytochrome oxidase.

It is of interest to note that most of these co-enzymes contain well-known vitamins of the B group as part of their molecules. Thus Co-enzymes I and II are derivatives of nicotinic acid, one of the B_2 vitamins, the flavine nucleotides contain riboflavine, one of the B_2 vitamins, and thiamine pyrophosphate is the pyrophosphate of vitamin B_1.

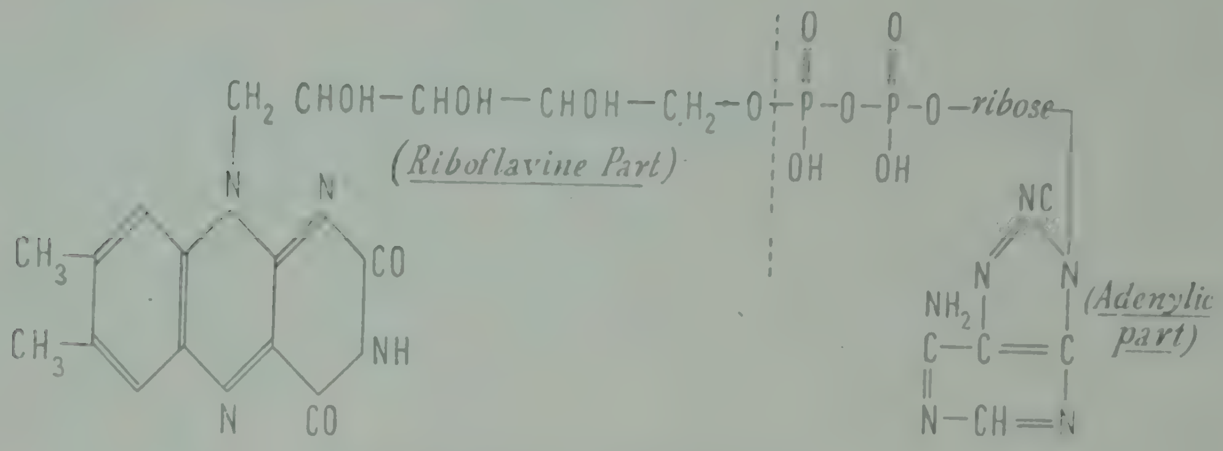
The Phospho-pyridino-nucleotides. There are two of these, viz. a Diphospho-pyridino nucleotide, called for short D.P.N. or "Co-enzyme I," and a Triphospho-pyridino-nucleotide, or T.P.N. or "Co-enzyme II." Of these Co-enzyme I, is probably more important in animal tissues. This co-enzyme is a compound of nicotinic amide, adenine, two molecules of d-ribose, and one of pyrophosphoric acid, and has the probable formula :—



Co-enzyme I (D.P.N.).

The compounds can also exist in a reduced form by introduction of t H- atoms near the N of the nicotinamide. For brevity it is customary to wri oxidised and reduced forms as D.P.N. and D.P.N.H_2, or Co.I and Co.I. respectively.

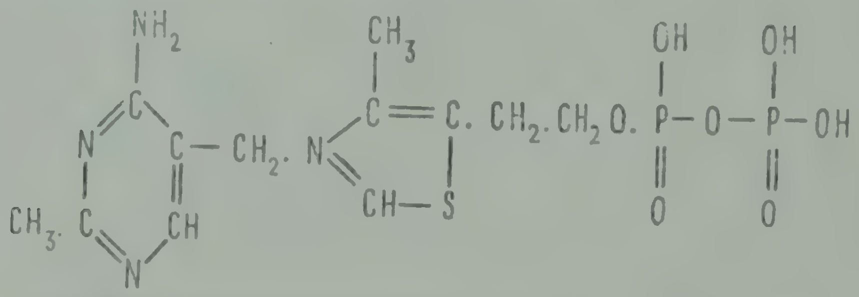
There are various *flavine nucleotides*, the dinucleotide having t structure :—



Flavinadenine dinucleotide.

This compound can also be reversibly reduced by taking on 2H into th rings of the riboflavine part of the molecule.

Thiamine pyrophosphate has the structure :—



Thiamine pyrophosphate (Co-carboxylase)

Enzymes as Synthetic Agents. Theoretically all enzyme actions are reversible, since the end-point represents an equilibrium in which free energy is at the minimum. In most cases, however, this equilibrium represents almost complete reaction in one direction, so that the energy of reversal is slight. But some enzymes have been found capable of catalysing reactions in the reverse direction and forming the original substrate, or, more often, a substance closely resembling it. This happens most readily if the concentration of normal end products is very great. Croft Hill * showed that if maltase were added to a 40 per cent. solution of glucose, a change took place in the reverse direction, which proceeded until 85 per cent. of the glucose was left. The sugar formed, a disaccharide, is not maltose, but the stereoisomeric sugar, iso-maltose.

In the same way it was found by Castle and Loewenhart that the hydrolysis of esters by lipase is a reversible reaction, the action of lipase being simply to hasten the attainment of the equilibrium point between the four substances—ester (or neutral fat), water, fatty acid and alcohol. It has been shown † that " plastein " resembling a protein can be resynthesised from protein hydrolysis products by the action of pepsin at a *p*H of 4·0 and at a relatively high temperature, *e.g.* 70° C. Addition of protein delays the synthesis. A similar synthesis is effected by trypsin at *p*H 5·7.

Perhaps the clearest example of enzymic synthesis is that of starch or glycogen. Peas and liver both contain a phosphorylase which can act on α-glucose-1-phosphate, to produce the polysaccharides starch and glycogen respectively.

If, as is believed, all enzyme actions are reversible, a possibility is opened of regarding the synthetic processes occurring in the living cell, as well as the processes of disintegration, as determined by the action of enzymes. It must be noted that actual reversal of a given enzyme action is only obtained with distinctness when dealing with concentrated solutions. The degree of synthesis, which would be produced in the very dilute solutions of glucose, etc., occurring in the animal cell, would therefore be infinitesimal. But if a mechanism were provided for the immediate removal of the synthetic product from the sphere of reaction, either by its being insoluble, by removing it to a different part of the cell or by building it up into some more complex body which was not acted on by the enzyme, the process of synthesis might go on indefinitely, and the infinitesimal quantities be summated to an appreciable amount. Further it should be noted that synthesis and breakdown of a given substance often occurs by different paths and under the influence of different enzymes.

---

* Croft Hill. *Trans. Chem. Soc.*, 1898, **73**, 634.

† Borsook, MacFadyean and Wasteneys. *J. Gen. Physiol.*, 1930, **13**, 295.

For Reference

KALCKAR (1941). The Nature of Energetic Coupling in Biological Syntheses. *Chem. Rev.*, **28**, 71.

NORTHROP (1939). The Chemistry of Pepsin, Trypsin and Bacteriophage. (New York: Columbia Univ. Press.)

SUMNER and SOMERS (1943). Chemistry and Methods of Enzymes. (New York: Academ. Press Inc.)

CLASSIFICATION OF ENZYMES

The names of enzymes are usually made by adding the suffix *-ase* to the name of the substrate, *e.g.* a peptidase acts on peptides: other names describe what the enzymes do, *e.g.* dehydrogenases remove hydrogen from their substrates. Several classifications have been proposed, but none of them are entirely satisfactory. A convenient provisional one, adapted from Baldwin, is given here, and illustrated by examples in the Tables below. Three main groups, each subdivided into sub-groups and classes are :—

- I. ***Splitting Enzymes***
  - (1) Hydrolases.
  - (2) Phosphorylases.
  - (3) Adding Enzymes.
- II. ***Transferring Enzymes***
  - (1) Oxidising and Reducing Enzymes.
  - (2) Other Transferring Enzymes.
- III. ***Isomerases***
  - (1) Simple Isomerases.
  - (2) Mutases.

EXAMPLES OF SOME IMPORTANT ENZYMES (after Baldwin)

I. SPLITTING ENZYMES

Group.	Class.	Example	Substrate	Products
(1) *Hydrolases*	Esterases	Lipase	Fat	Fatty acids and glycerol.
,,	,,	Phosphatase	Phosphoric esters	H_3PO_4 and alcohol.
,,	Carbohydrases	Amylase	Starch	Maltose.
,,	,,	Maltase	Maltose	Glucose.
,,	Peptidases (Endo-peptidases).	Pepsin	Proteins	Proteoses and peptones.
,,	,,	Trypsin	Proteins	Peptones and amino-acids.
,,	Exo-peptidases	Carboxypeptidase	Peptides	Amino-acids.
,,	,,	Dipeptidase	Dipeptides	Amino-acids.
,,	Deaminases	Adenase	Adenine	Hypoxanthin $+ NH_3$
,,	Deamidases	Urease	Urea	$CO_2 + NH_3$
(2) *Phosphorylases*	Amylophosphorylases	Liver phosphorylase	Glycogen and phosphate.	glucose-1-phosphate.
(3) *Adding enzymes*		Catalase	H_2O_2	$H_2O + O_2$
,,		Carbonic anhydrase	H_2CO_3	$H_2O + CO_2$

II. TRANSFERRING ENZYMES

Group 1. Oxido-reducing Enzymes (Transfer 2H)

Class	Example	H-donator[*]	H-acceptor
Anaerobic dehydrogenases	Succinic deh.	Succinate	Cytochrome *c* (or MB
,, ,,	Lactic deh.	l-lactate	Co I.
Aerobic dehydrogenases	Amino-acid oxidase	Amino-acids	O_2 (or MB).
	Xanthine oxidase	Xanthine, etc.	O_2 (or MB).
Flavoprotein	Cytochrome reductase	Reduced Co II	Cytochrome *c* (or MB
Aerobic oxidases	Cytochrome oxidase	Cytochrome *c*	O_2.
,, ,,	Tyrosinase	Phenols	O_2
Peroxidases	Cytochrome peroxidase	Reduced Cyt. *c*	H_2O_2.

Group 2. Other Transferring Enzymes

Group transferred	Class	Example	Group donator	Group acceptor
H_2PO_3—	Transphosphatases	Hexokinase	A.T.P.	Glucose.
,,	,,	Lohmann enzyme	A.T.P.	Creatine.
—NH_2	Transaminases	Glutamic Tr-am	Glutamic acid	α-Keto acids.
—CH_3	Transmethylases	Unnamed	Methionine	Ethanolamine.

III. ISOMERASES

	Example	Reaction
Group 1. Simple isomerases	Oxoisomerase	Glucose 6-phosphate ⇌ fructofuranose 6-phosphate.
Group 2. Mutases	Phosphoglucomutase	Glucose 6-phosphate ⇌ glucose-1-phosphate.

I. Splitting Enzymes. These catalyse the breaking up of a larger molecule into two or more smaller ones.

(1) HYDROLASES. These enzymes bring about *hydrolysis* of the larger molecules on which they act, by the general reaction :—

$$RR' + HOH \rightleftharpoons ROH + R'H$$

Thus, the enzyme maltase catalyses the hydrolytic breakdown of maltose into two molecules of glucose

$$C_{12}H_{22}O_{11} + H_2O = 2C_6H_{12}O_6$$

The digestive enzymes acting on carbohydrates, proteins and fats all belong to this class.

The reverse change, that of dehydration, is also effected apparently with equal facility by the living cell ; thus the amino-acids are built up, first into polypeptides, and these again into the complex proteins of the cell.

Besides the reactions in which there is a difference in the amount of free water on the two sides of the equation, it seems probable that hydrolysis and simultaneous dehydration at different parts of the molecule determine a number of chemical transformations, which at first sight seem to involve a simple splitting of the molecule.

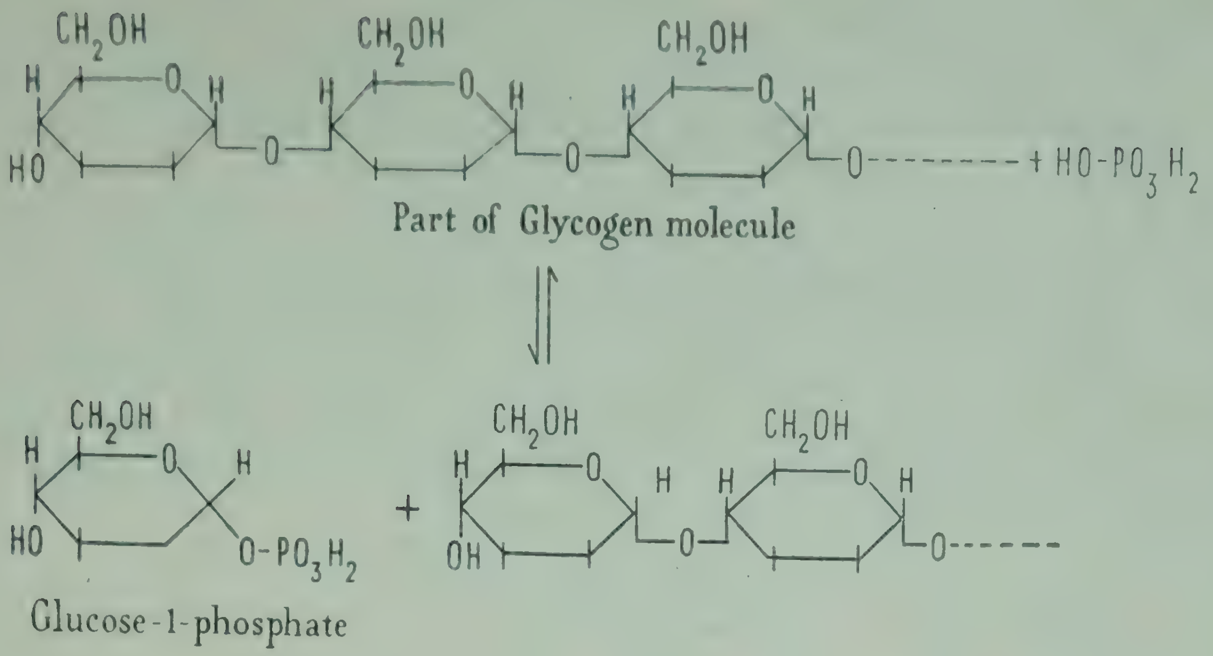
(2) PHOSPHORYLASES. With these enzymes the larger molecule is broken up by reaction with phosphoric acid, thus :—

$$RR' + H\,\vdots\,OPO_3H_2 \rightleftharpoons ROPO_3H_2 + HR'$$

so that a phosphoric group is introduced into one of the products. Thus a phosphoryl group can be introduced from phosphate into glucose, cut off from a glycogen molecule.

---

[*] MB. = Methylene blue.



(3) ADDING ENZYMES. Merely split a molecule into two parts,

$$RR' \rightleftharpoons R + R'$$

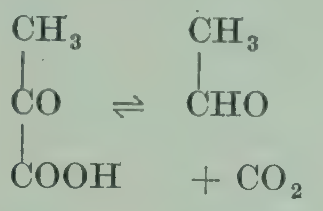
or, when acting in reverse, cause the two smaller parts to unite to form the larger molecule. Thus the simplest case is that of the enzyme *catalase*, which splits hydrogen peroxide into water and oxygen (but not reversibly) :

$$2H_2O_2 \longrightarrow 2H_2O + O_2$$

Another example is *carbonic anhydrase* which catalyses the conversion :

$$CO_2 + H_2O \rightleftharpoons H_2CO_3$$

and another is *carboxylase*, which converts pyruvic acid into acetaldehyde and CO_2 :—



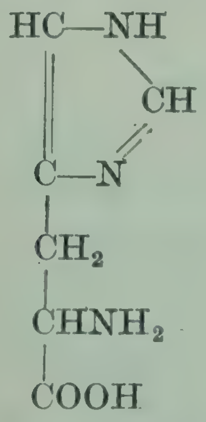
$$\begin{array}{l} CH_3 \\ | \\ CO \\ | \\ COOH \end{array} \rightleftharpoons \begin{array}{l} CH_3 \\ | \\ CHO \\ \\ + CO_2 \end{array}$$

This reaction is one stage in the formation of alcohol from glucose by yeast ; pyruvic acid is formed at one stage, and is then acted upon by carboxylase, of the yeast, with its co-enzyme, forming acetaldehyde and CO_2. The acetaldehyde is then further reduced to alcohol by other enzymes.

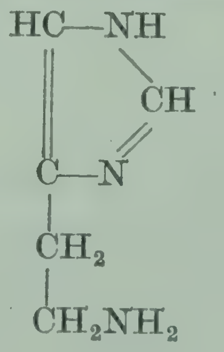
Many amino-acids when subjected to the action of bacteria lose a molecule of carbon dioxide and are converted into a corresponding amine. There is a large number of enzymes capable of decarboxylating amino-acids in this way : they are called the *amino-acid decarboxylases*, and each one is specific for a particular amino-acid.

For instance lysine is converted into pentamethylene diamine or cadaverine ;

and Histidine



becomes Histamine



In the same way ornithine is converted into tetra-methylene diamine or putrescine. Other examples of this process of decarboxylation are :

β-phenylethylamine $C_6H_5.CH_2.CH_2.NH_2$, from phenylalanine.
Para-hydroxyphenylethylamine $HO.C_6H_4.CH_2.CH_2.NH_2$, from tyrosine.

It is worthy of note that all the changes which we have so far been considering involve but little expenditure of energy.

II. **Transferring Enzymes.** These enzymes catalyse reactions i which a group is detached from one molecule, called the *donator* molecule, an transferred to another molecule, called the *acceptor* molecule, according to t general reaction :—

$$RA + R' \rightleftharpoons R + R'A$$

This is a large and important group, and includes enzymes concerned wit biological oxidations and reductions.

(1) OXIDISING AND REDUCING ENZYMES. Oxidation and reduction ar strictly speaking, the result of a transfer of electrons, oxidation being the los of an electron or of electrons, and reduction the gain. Most frequently oxidatio consists chemically in either a gain of oxygen, *e.g.*,

$$RCHO \longrightarrow RCOOH$$

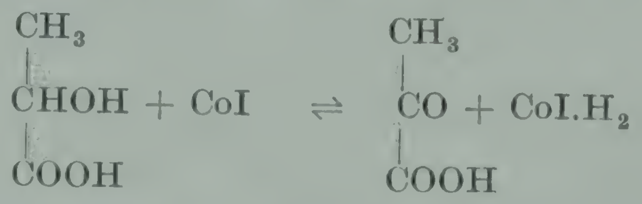
or loss of hydrogen atoms, *e.g.*,

$$RCH_2OH \longrightarrow RCHO$$

by the molecule under consideration, but whichever process happens, it i important to remember that two molecules are involved in the process, on being oxidised and the other reduced ; *oxidation of any substance is alway accompanied by reduction of some other substance.* We may conveniently regar oxidation as being most often a transference of H- atoms from the substanc oxidised to the substance reduced, and less often as a transference of oxyge in the opposite direction.

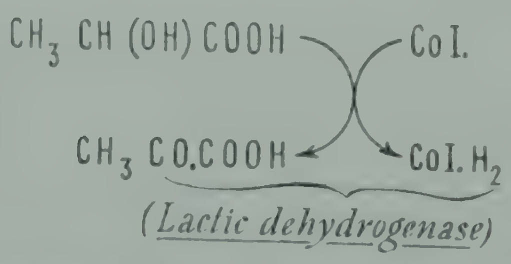
The oxidising and reducing enzymes thus are enzymes which transfer the group 2H : the substance losing the 2H is thereby oxidised and is called the *hydrogen donator*, while the substance gaining the 2H is simultaneously reduced, and is called the *hydrogen acceptor*.

In these oxidation-reduction reactions, various *co-enzymes* often play an important part by acting alternatively as hydrogen acceptors or donators. For example, the enzyme lactic dehydrogenase removes 2H from lactic acid (the donor) and so converts it into pyruvic acid, and simultaneously reduces Co-enzyme I (the acceptor) by transferring the 2H to it :—



$$\begin{matrix} CH_3 \\ | \\ CHOH \\ | \\ COOH \end{matrix} + CoI \rightleftharpoons \begin{matrix} CH_3 \\ | \\ CO \\ | \\ COOH \end{matrix} + CoI.H_2$$

This reaction can conveniently be shown by a scheme of the type introduced by Baldwin :—

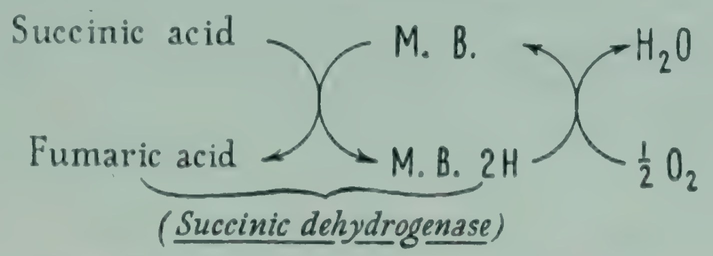


As we shall see later, the reduced CoI can subsequently be reoxidised by other reactions, and so can oxidise another molecule of lactic acid, and so on.

In some instances the 2H is ultimately transferred to atmospheric oxygen, which is reduced to H_2O or H_2O_2. Such oxidations by free oxygen are usually the result of a chain of reactions in which a series of different enzymes and co-enzymes are concerned, but are clearly of great interest since they provide a clue to the way in which oxygen is used by living cells to effect oxidations. As a simple illustration of the kind of thing that happens we may choose the action of *succinic dehydrogenase*. This enzyme can transfer hydrogen to methylene blue, which acts in place of a normal tissue H-acceptor. If washed minced muscle be placed in a dilute solution of methylene blue, a little succinic acid added to it, and the mixture placed out of contact with oxygen, and kept warm, the methylene blue is quickly reduced to a colourless compound, and an equivalent amount of succinic acid is oxidised to fumaric acid. The reaction is one of hydrogen transfer, the succinic acid losing two atoms :

$$C_4H_6O_4 \rightarrow C_4H_4O_4 + 2H,$$

and the 2H is taken up by the methylene blue. On now shaking the mixture with air, oxygen is taken up, and the colour is restored, owing to the reformation of methylene blue. In presence of oxygen, therefore, the oxidation will go on continuously to completion, as in the scheme :—

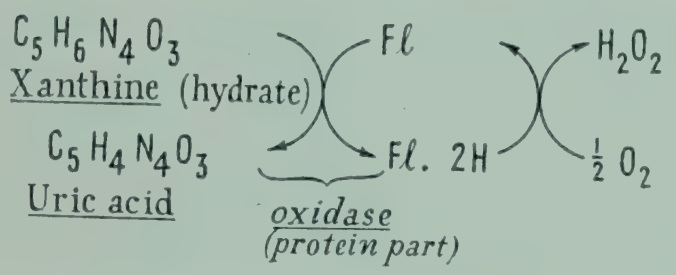


The direction of the arrows shows how the methylene blue is alternatively reduced and oxidised.

Some enzymes seem at first sight to be able to bring about a direct oxidation by free oxygen, with formation of hydrogen peroxide or water ; such enzymes are called *oxidases*. Thus the enzyme *xanthine oxidase* converts xanthine into uric acid :—

$$\underset{\textit{Xanthine}\text{ (hydrate)}}{C_5H_5N_4O_2OH} + O_2 = \underset{\textit{Uric acid}}{C_5H_4N_4O_3} + H_2O_2$$

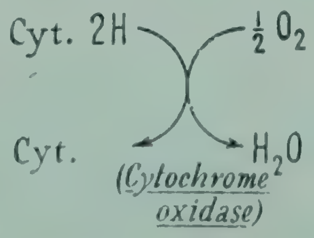
with formation of hydrogen peroxide. But the apparently direct oxidation is in reality a stagewise change. The xanthine oxidase is a conjugated protein, consisting of a protein part, the "real" oxidase enzyme, and a prosthetic group, flavine adenine dinucleotide (Fl) which is really a reversible hydrogen acceptor co-enzyme permanently attached to the protein. When reduced it can be reoxidised by free oxygen. Hence the whole reaction is as in the scheme :—



Any hydrogen peroxide formed in such reactions in the tissues is at once removed by the enzyme *catalase*, by which it is broken down to water and oxygen.

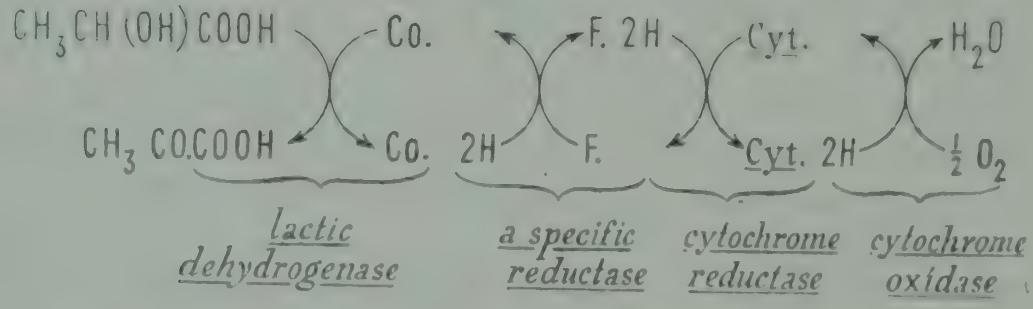
Aerobic Oxidases. There are, however, some enzymes which seem able to use free oxygen directly as a H- acceptor, the H- donator being thereby oxidised. One of the most important of these enzymes is called *cytochrome oxidase* : its H- donator being reduced cytochrome -*c*.

The cytochromes, of which there are several, *a*, *b*, *c*, etc., are conjugated proteins containing an iron-porphyrin group, which are present in all living cells capable of using oxygen. The iron can exist in the ferri- or ferro- form according as the state of the cytochrome is oxidised or reduced. Now, cytochrome oxidase in presence of oxygen *oxidises* reduced cytochrome -*c* to the oxidised form and reduces the oxygen to water :—



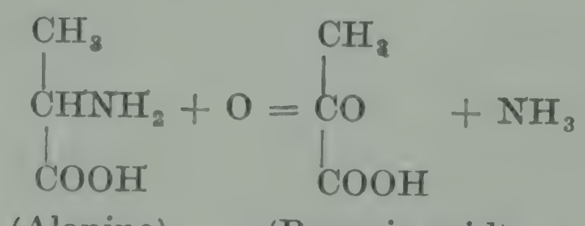
This action is prevented by the presence of traces of HCN, which paralyses the enzyme : it is probably on this account that cyanides suppress the oxidative processes of the tissues and are poisonous.

This oxidation of cytochrome -*c* by its oxidase is, however, the termina reaction in an oxidative chain: the cytochrome has previously been reduced by acting as H- acceptor for a dehydrogenase, and this dehydrogenase contain flavinadenine mono-nucleotide as its prosthetic group. It is called *cytochrom reductase*; its action is inhibited by narcotics like urethane or alcohol. Thi enzyme *reduces* cytochrome -*c* and oxidises reduced CoII, but there are probably other cytochrome reductases that can reduce cytochrome -*c* and oxidis reduced CoI. If so, we can expect that the complete series of reactions that lead to oxidation with ultimate oxygen absorption to be somewhat as follows (writing Co for co-enzyme I (or II), F for the flavine nucleotide and *Cyt.* for cytochrome), and taking as an example the oxidation of lactic acid :—



Deamination. This process involves the splitting off from an amino-acid of an NH_2 group as ammonia. Many tissues of the body appear to have this power: if, for instance, to slices of surviving liver or kidney, an amino-acid be added, ammonia is set free in proportion to the amount of amino-acid present.

Although the simplest mode of deamination of an amino-acid would apparently be the hydrolytic one, with formation of ammonia and a hydroxy acid, work by Neubauer and by Krebs showed that deamination is accompanied in the first place by oxidation, so that the first intermediate product formed is an α-ketonic acid.

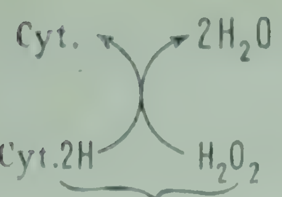


$$\begin{array}{l}CH_3\\ |\\ CHNH_2\\ |\\ COOH\end{array} + O = \begin{array}{l}CH_3\\ |\\ CO\\ |\\ COOH\end{array} + NH_3$$

(Alanine) (Pyruvic acid)

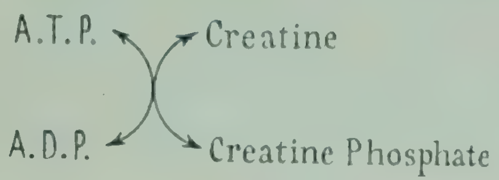
Little is known of the enzymes which effect this change, but one of them at least, viz, *l*-glutamic dehydrogenase functions with co-enzymes I or II, and possibly the others are similar though less specific as regards their substrates.

Peroxidases. It has long been known that hydrogen peroxide is able to effect oxidative changes similar to those observed in the animal body. Thus Dakin showed that the series of fatty acids, under the action of hydrogen peroxide, undergoes progressive oxidation in the β position, so that the carbon atoms are removed from the long chain of fatty acids two at a time. The action of hydrogen peroxide is much quickened by the presence of a trace of iron in the ferrous condition (Fenton's reaction): the iron thus acts as a catalyser. As we have seen above, hydrogen peroxide is sometimes formed as a result of the action of some of the oxidases. This is removed usually by the enzyme *catalase*, but it was at one time supposed that important oxidative changes were brought about by the activation of the peroxide by an enzyme called *peroxidase*. Such enzymes do exist, but are more readily demonstrable in plant than in animal tissues. They are detected by the fact that they oxidise certain substances to form coloured products. Thus guaiacum resin or guaiaconic acid is turned blue in presence of a peroxide and a peroxidase; if the enzyme preparation also generates H_2O_2 by the action of other enzymes, then it turns guaiacum blue without addition of a peroxide. The only important animal peroxidase appears to be *cytochrome peroxidase*, which oxidises reduced cytochrome -*c* by a reaction alternative to that by cytochrome oxidase.



(*cytochrome peroxidase*)

(2) Other Transferring Enzymes. These enzymes catalyse the transfer of various other groups from a donator to an acceptor molecule. The most important transferences of groups are those of the phosphoric acid group, by *Transphosphatases*, of the amino group by *Transaminases*, of the methyl group by *Transmethylases*, etc. For instance, one of the phosphokinases, transfers a phosphoric group, with energy-rich bond intact, from adenosine triphosphate to creatine :—



III. **Isomerases.** These enzymes cause their substrates to change into isomeric forms. For instance *oxoisomerase* catalyses the change of glucose 6-phosphate into fructose 6-phosphate, by molecular rearrangement. The *Mutases* are isomerases which cause more extensive internal rearrangements.

Oxidation–reduction Potentials. Although for most purposes it is sufficient to consider the phenomena of oxidation and reduction as simple chemical reactions, it is often valuable to regard them from the physico-chemical standpoint, as questions of transfer of electrons. In this sense oxidation is regarded as the loss of an electron, or of electrons, from an atom, and reduction as a gain of electrons. Thus when ferrous iron is oxidised to ferric iron an electron or charge of negative electricity is lost, so that the ferric iron atom gains a positive charge. This is expressed by the electro-chemical equation :—

$$\mathrm{Fe}\ {+}{+} \rightleftarrows \mathrm{Fe}\ {+}{+}{+} + e,$$

the quantum *e* for 1 gramme-atom being 1 Faraday, or 96,500 coulombs of negative electricity. It will be seen that this conception is really more accurate than one involving a transfer of either hydrogen or oxygen, since it covers all cases, *e.g.* the oxidation of ferrous chloride to ferric chloride, as above. Oxidation is always accompanied by reduction in some other substance, which thereby gains the electron lost by the oxidised substance. In cases where there is merely a transfer of electrons in this manner there is of course no energy change, but often there is liberation of some of the energy as heat or in some other form.

In order for oxidation to occur, there must be a readiness for the electron to escape, and the greater its escaping tendency, or *fugacity*, the more readily will oxidation take place. This tendency can be measured for any dissolved substance by the determination of the *oxidation-reduction potential*, which may be explained as follows : Suppose that into a readily oxidisable solution, *e.g.* $FeCl_2$, we dip a clean gold surface, then the Fe electrons, by their tendency to escape, will give to the electrode a negative charge, which will be greater the more the electron-fugacity of the substance. We thus have a half-cell, which may be combined with another half-cell, *e.g.* a hydrogen electrode, very much as described on p. 65, to form a cell the potential of which can be measured by some method which will not draw much current, *e.g.* an electrometer. The negativity of the gold electrode, as thus measured, is a measure of the ease of oxidisability of the substance into which it is dipped, or alternatively, of the reducing power of that substance towards other substances. Conversely, the less negative (or more positive) the electrode the greater the reducibility (or oxidising power) of the substance.

It is a convention to express the results in terms of what is called E_h, which is the reading, in volts, of the oxidation-reduction cell when set up with a normal hydrogen electrode, *i.e.* N.HCl. It follows from what has been said above that a solution of a substance capable of *reversible* oxidation and reduction will oxidise substances whose solutions have a more negative E_h (*i.e.* it will take up electrons from them), and reduce substances which have a more positive E_h than its own (*i.e.* it will give up electrons to them). Thus the E_h of the Ox/Red. flavine enzyme is $-$ 0·063 volts and that of Ox/Red. coenzyme I is $-$ 0·29, *i.e.* the flavin system can oxidise the $CoIH_2$. By determination of the E_h of a given oxidation-reduction system also, we can calculate the *free energy* of the change. For instance, it can be shown that for the reaction :—

$$\text{lactate} \rightleftharpoons \text{pyruvate} + 2H^+ + 2e$$
$$\text{the } \triangle F = -\ 8{\cdot}304 \text{ Cal.}$$

If we take a mixture of such a substance in the oxidised and reduced forms, it wil have a more negative E_h the more of the reduced form it contains, and *vice versa.* Many dyes are capable of oxidation and reduction, and since the reduced forms are often colourless, such dyes, when added to mixtures of other substances, become oxidised or reduced according to the oxidising or reducing potentials of the whole system in which they are present. We can use these dyes to assess the oxidation-reduction potentials in a manner somewhat analogous to the use of coloured buffer-dyes for the assessment of the hydrogen ion concentration of a mixture.

It has long been known that dyes such as methylene blue can often be reduced by most tissues; others, such as indophenol blue,* are reduced by some tissues but not by others, such as heart or brain: alizarin blue has so high a reduction potential that it could only be reduced by a few organs, *e.g.* lung, liver. Further, we know that such reversibly oxidisable substances, by being alternately reduced and oxidised, can act as hydrogen or oxygen carriers, and thus act with oxidation or reduction catalysts. Many naturally occurring substances, both in plants and animals, have this property, *e.g.* glutathione, ascorbic acid, various flavins, etc. Such substances, loosely linked to protein, form the oxidising enyzmes, and "when a group of these enzymes and their oxidisable substances act together . . . the oxidation reactions proceed, as a rule, in a series, the system of more negative potential being oxidised by the reduction of the system of more positive potential, and so on, in a series of graded steps up to the reduction of molecular oxygen. So, like locks in a canal, these reversible systems release gradually the oxidation energy" (Barron).

For Reference

Barron (1939). Cellular Oxidation Systems. *Physiol. Rev.*, **19**, 184.

Baumann and Stare (1939). Co-enzymes. *Physiol. Rev.*, **19**, 353.

Clark, W. Mansfield (1925). Recent Studies on Reversible Oxidation-Reduction in Organic Systems. *Chem. Rev.*, 1925, **2**, 127.

Davenport (1946). Carbonic Anhydrase in Tissues other than Blood. *Physiol. Rev.*, **26**, 560.

Green (1940). Mechanisms of Biological Oxidations. (London: Cambridge Univ. Press.)

Hellerman (1937). Reversible Inactivation of Certain Hydrolytic Enzymes. *Physiol. Rev.*, **17**, 454.

Meyerhof *et al.* (1942). A Symposium on Respiratory Enzymes. (Univ. Wisconsin Press.)

Northrop (1939). Crystalline Enzymes. (New York: Columbia Univ. Press.)

Northrop (1937). The Formation of Enzymes. *Physiol. Rev.*, **17**, 144.

Ogston and Smithies (1948). Some Thermodynamic and Kinetic Aspects of Metabolic Phosphorylation. *Physiol. Rev.*, **28**, 283.

FOR GENERAL REFERENCE OF SUBJECT-MATTER OF BOOK I

Baldwin (1947). Dynamic Aspects of Biochemistry. (Camb. Univ. Press.)

Bayliss (1924). Principles of General Physiology. (Longmans, Green.)

Cowdry, E. V. (1924). General Cytology. (Chicago.)

Glasser (1944). Medical Physics. (Year Book Publishers, Chicago).

Gray, J. (1931). A Text-book of Experimental Cytology. (Camb. Univ. Press.)

Harrow (1946). Text-book of Biochemistry (W. B. Saunders Co.).

Heilbrunn (1937). An Outline of General Physiology. (Saunders.)

Höber (1945). Physical Chemistry of Cells and Tissues. (Churchill.)

Wishart (1931). Groundwork of Biophysics. (London: Bell & Sons.)

---

* Ehrlich. "Das Sauerstoffbedürfniss des Organismus," Berlin, 1885.

BOOK II

TISSUES SUBSERVING MOVEMENT AND CONDUCTION—MUSCLE AND NERVE

HISTORICAL NOTE

Anatomical dissection informed the ancients of the approximate uses of the various muscles, and the injuries which men sustained in the pursuit of war and assassination doubtless revealed to them the essential part played by the nervous system in the performance of voluntary movements. Galen (a.d. 130–200) recognised the involuntary muscles, realised the essential action of antagonistic muscles and of muscular tone, and showed by experiment that section or ligation of a motor nerve resulted in paralysis of its muscle. According to his speculations, which were an extension of those made some 500 years earlier by Erasistratus of the Alexandrian school, the working of the muscles was attributed to the passage into them from the nerves, of a pneuma, or spirit, the so-called psychic or animal spirit, which was engendered by the central nervous system. Although the succeeding fourteen centuries saw, towards its latter end, an extension of anatomical knowledge, there seems to have been little departure, and indeed little chance for a departure, from the Galenical outlook upon function. Even after the introduction of the microscope by Galileo in 1609, it was over half a century before the muscle fibre was described in some detail by Stensen (1664) and, more fully, by Leeuwenhoek (1715). It was not until, in the 17th century, the science of Physics, and later, that of Chemistry, began to take shape, that an outlook on biological science comparable with that which obtains to-day, became possible. Hence, with slight variations, the Galenical view, that in contraction the muscle became *distended* with the animal spirits, runs as a thread through the writings of 1500 years. Descartes (1662), Mayow (1674), and Borelli (1680) all subscribed to it in one form or another. Stensen showed that after ligation of its artery a muscle lost its power of contracting. Mayow's view was that the animal spirits consisted of " nitro-aerial spirit " (oxygen), and that, in the muscle, these meet with " salino-sulphureous particles " (combustible material), carried thither by the blood, with the result that there is an " effervescence " in the muscle causing it to swell up. Hence, in exercise, there is deeper breathing, to supply more " nitro-ærial spirit " : hence also, ligation of the nerve, by obstructing the flow of " nitro-aerial spirit," or of the artery, by obstructing the supply of salino-sulphureous particles, deprived the muscle of the power to act. The proof that, in fact, the volume of the muscle is not appreciably changed on contraction was furnished by Jan Swammerdam towards the end of that century, but was not published till 1738. He was the first to use the frog muscle-nerve preparation. Francis Glisson, about 1677 had also shown, by the use of a simple plethysmograph, that a limb shows no increase in volume, but rather a decrease, when its muscles contract : he also introduced the idea of excitability, or, as he called it, " irritability " though it was not adopted generally for nearly a century.

Until about 1664, or even later, doubt existed as to which part of a muscle was responsible for its contraction, a prevalent view being that it was the connective tissue part : simultaneously in 1664, William Croone and Stensen decided in favour of the fleshy part as the active tissue.

Early in the 18th century, the first text-book of physiology, Boerhaave's *Institutiones medicæ* (1708), marked the commencement of formal treatment of the subject, and later Albrecht von Haller's *Elementa physiologiæ corporis humani* (1757–65) collected together all the contemporary knowledge. It contained the results of the author's own work on muscle and nerve, and on the irritability of tissues. He studied the effects of various stimuli on the muscles and nerves of live and of recently killed animals, and concluded that the animal spirits of the older writers was, in fact, a fluid " unlike everything else," and was passed along the nerves, which were tubular in structure. The term " excitability " was substituted for irritability by John Brown in 1780.

In 1743 the mathematician Hausen, with knowledge of the electrical machine, put forward the suggestion that nervous action was electrical in nature, but Haller, the then dominant physiologist, rejected this view on the grounds that ligation of the nerve, though incapable of stopping the passage of an electric current, suspended the propagation of the nervous action.

The invention of the Leyden jar about the middle of the 18th century soon led to the discovery that electricity supplied an admirable stimulus for the excitable tissues, and towards the close of the century, in 1786, Galvani (publ. anonymously in 1794) made th original observations from which sprang the knowledge of animal electricity, and adum brated the hypothesis of the essentially electrical nature of the conducted process in nerve a view which the introduction of the telegraph in 1835 further encouraged.

During the 19th century, events and discoveries followed at an accelerating pace, t culminate in an avalanche of fact and theory in the present century. The period of mos fruitful harvest was on the virgin soil between 1835 and about 1880, and a mere catalogu of salient events must suffice here.

The study of the heat production of contracting muscle by means of thermo-electri couples or thermopiles was due to Bequerel and Breschet (1835), Helmholtz (1864 Heidenhain (1864), and Fick (1882), the last two considering the phenomena from th thermodynamical standpoint. The researches of A. V. Hill for decades from 1910 adde largely to knowledge in this field.

Animal electricity became the subject of more exact study after the invention b Nobili, in 1827, of the astatic couple, and Matteucci and Du Bois Reymond made mos of the pioneer discoveries in this field between 1840 and 1843.

The main laws of the excitation of nerve by the constant current had been sketche out by about 1848. The way for further investigations on the nerve impulse was opene by the introduction into physiology of the induction coil by Du Bois Reymond in 1849 and the time relations of muscular contraction were first worked out by Helmholtz in 1850. Du Bois Reymond about 1849 accepted the electrical changes in nerve as a sig of the transit of the conducted disturbance, and in 1859 he added improvements to the non-polarisable electrodes introduced in 1854 by Regnauld.

Views concerning the chemical events associated with muscular contraction developed much later. The increase of oxygen utilisation during muscular exercise observed in 1789 by Lavoisier and Seguin, were more exactly confirmed by Regnault and Reiset in 1849. The formation of lactic acid in contraction, discovered by Berzelius in 1841, was stressed by Hermann in 1867. It was Hermann also who, with Pflüger, postulated the existence in muscle of a labile component, "*inogen*," the breakdown of which provided the energy for contraction. Glycogen was discovered by Claude Bernard in 1857, Creatine by Chevreuil in 1835, and Creatine phosphate by the Eggletons and independently by Fiske and Subarrow in 1928. The construction of plausible chemical hypotheses regarding muscular contraction began with the work of Fletcher and Hopkins in 1907, and is still in progress. The discovery in 1928 of adenosine triphosphate in muscle, by Lohmann, has turned out to be one of the most important discoveries in biological chemistry. Its properties strangely recall those of Hermann's inogen.

The study of the molecular structure of muscle protein by X-ray spectra by Meyer and Astbury, from 1929 onwards, opened up a new line of enquiry into the mechanism of muscular contraction.

An idea of the rapid development of the subject in the mid-19th century is given by considering the views on the velocity of the nervous impulse. In 1844 the great Johannes Müller stated that it must be similar to the velocity of light, and could probably never be measured: reflexes, similarly, were regarded as instantaneous. But in 1850 Helmholtz, by a procedure closely resembling that now carried out in student exercises, measured the velocity of the transmission in frog's nerve, and he also measured a reflex time in 1854.

The electrotonic currents of nerve were demonstrated by Du Bois Reymond in 1843, and the whole subject of electrotonus was extensively studied by Pflüger, whose book on the subject appeared in 1859. The "core model" illustrating its phenomena (Matteucci, 1863) was developed by Hermann (1873), and was a source of much controversy.

The action of curare was studied by Claude Bernard in 1857; by its use he showed that muscle was independently excitable, and that a chemical substance could be used as an implement of physiological research and analysis.

The regeneration of divided nerves was first observed by Cruickshank in 1776: the process of degeneration was studied by Waller in 1850, and shown to be of use for tracing the cells of origin of the divided nerves.

The relative unfatigability of nerve was first properly demonstrated by Bernstein in 1877. Various attempts were made to show that nerves generated heat when impulses passed along them, but the most accomplished experimenters of the 19th century, Helmholtz (1848) and Heidenhain (1868), failed, where others claimed to have succeeded. It was not until 1931 that a heat production was demonstrated in nerve fibres.

The all-or-none character of nerve-fibre transmission was first recognised about 1909 by the work of Keith Lucas.

In the last two or three decades, refinements of physical and chemical methods have permitted the investigation in detail of many aspects of the physiology of muscle and nerve that were previously not amenable to such precise examination. Among these advances in technique are the employment (about 1921) of radio amplification together with the use of the cathode ray oscillograph, and the adaptation of many micro-chemical methods. The idea of the humoral transmitters grew out of Otto Loewi's demonstration in 1921 that acetyl-choline was formed in the heart on vagus stimulation.

CHAPTER X

VOLUNTARY STRIATED MUSCLE

MOVEMENT and conductivity are of the very essence of life phenomena. The potentiality or property of exhibiting movement, macroscopic or microscopic, is probably inherent in every living cell; so also is the potentiality or property of conducting a state of change (often excitatory) from one part of the cell to another. In all the higher animals the power of movement is a condition for survival, and they accordingly possess tissues in which the power of movement is especially well developed. These are the muscular tissues. Further, speedy correlation between the various parts of the body is effected by tissues of highly developed conductivity for states of excitation. These are the nervous tissues.

In the higher animals we may distinguish three varieties of muscle. In all three varieties, movements are effected as a result of a shortening, or *contraction*, of a fibre or cell, usually of elongated form. All movements that require to be sharply carried out are effected by means of striated muscular tissue and, as these movements are often under the control of the will, the muscles are generally spoken of as *voluntary*. Whether controlled by the will or not, however, they are normally always caused to contract as the result of excitation received by way of a nerve, *i.e.* they are *nerve-operated*. If its nerves are severed, the muscle is useless and soon degenerates. Plain or *involuntary* muscles form sheets or closed tubes surrounding the viscera, the blood vessels, etc. By their slow prolonged contractions they serve to maintain and regulate the flow of the contents of these organs. Intermediate in properties as well as in structure between these two classes is the *heart or cardiac muscle*.

Plain and cardiac muscle can and do normally contract independently of nerves, though nerves can alter these contractions, *i.e.* they are *nerve-regulated* but not nerve-operated.

The properties of contractile tissues have been most fully investigated in the voluntary muscles, particularly those of the frog. The selection of cold-blooded animals is guided by the fact that it is possible to isolate the muscle from the rest of the body and to keep it alive long enough to study its reactions during a considerable time. We shall therefore deal at length with the properties of skeletal muscle, pointing out incidentally in what respects heart muscle and involuntary muscle differ from it.

Much of our knowledge on the subject of muscle has been derived from the study of the isolated gastrocnemius and sartorius muscles of the frog. The gastrocnemius, with the attached sciatic nerve, can conveniently be removed from the body, and is called a *nerve-muscle preparation*. On account of a penniform arrangement of the muscular fibres, while the actual shortening of the gastrocnemius is but small, its strength of contraction is considerable.

The sartorius muscle is of especial value on account of the fact that its fibres are parallel. When a greater mass of approximately parallel fibres is necessary, recourse may be had to the gracilis and semimembranosus muscles together.

Other muscles in the frog used for particular purposes are the mylohyoid and the dorsocutaneous muscles. The mylohyoid muscle, which lies on the central surface of the tongue, has the advantage that its fibres lie in close contact with a lymph space occupying the centre of the tongue. If any drug be injected into this lymph space it acts with extreme rapidity on the muscle fibres, so that this preparation is useful for the study of the action of different substances on muscle fibres.

Experiments on isolated muscle fibres have often been carried out. They are valued in fundamental observations because the results obtained are supposed to be unequivocal.

Present-day investigations employ, to an increasing extent, mammalian, or human, muscle as their material.

STRUCTURE OF STRIATED MUSCLE

The striated muscles form up to 50 per cent. by weight of the body. Eac muscle is embedded in a layer of connective tissue and is made up of a aggregation of muscular fibres 10 to 100μ in diameter, which are united int bundles by means of connective tissue. The individual fibres vary much i length and may exceed 10 cm., though often each fibre is shorter than th bundle of which it forms a part; most of them taper towards one or both end At each end of the muscle the fibres are firmly united by their blunt ends t the tendon of the muscles which are attached to bones. The fine taperin ends of the fibres end in the connective tissue between the fibres. Runnin in the connective tissue framework of the muscle we find a number of bloo vessels, capillaries and nerves.

The living muscle fibre exhibits a series of alternate light and dark band arranged at right angles to its long axis and enclosed in a structureless laye the sarcoplasmatic membrane. The whole fibre is encased in a tough sheath the sarcolemma (Fig. 32). Lying under the sarcolemma are a number of oval nuclei embedded in granular protoplasm. Each fibre may be artificially split into a bundle of parallel fibrils, which are embedded in an interstitial material called sarcoplasm. Each fibril consists of a column of compartments (after the fashion of a minute stick of bamboo), 2 to 3μ in length and each called a *sarcomere*. Each sarcomere is more transparent at the two ends than in the middle, thus giving rise to the appearance of light and dark striæ. In a fibre, corresponding structures in adjacent fibrils are side by side, thus giving rise to the cross striation. In the middle of the light band is a line or row of dots (often appearing double) called Krause's membrane or the Z line: this is the boundary of the sarcomere.

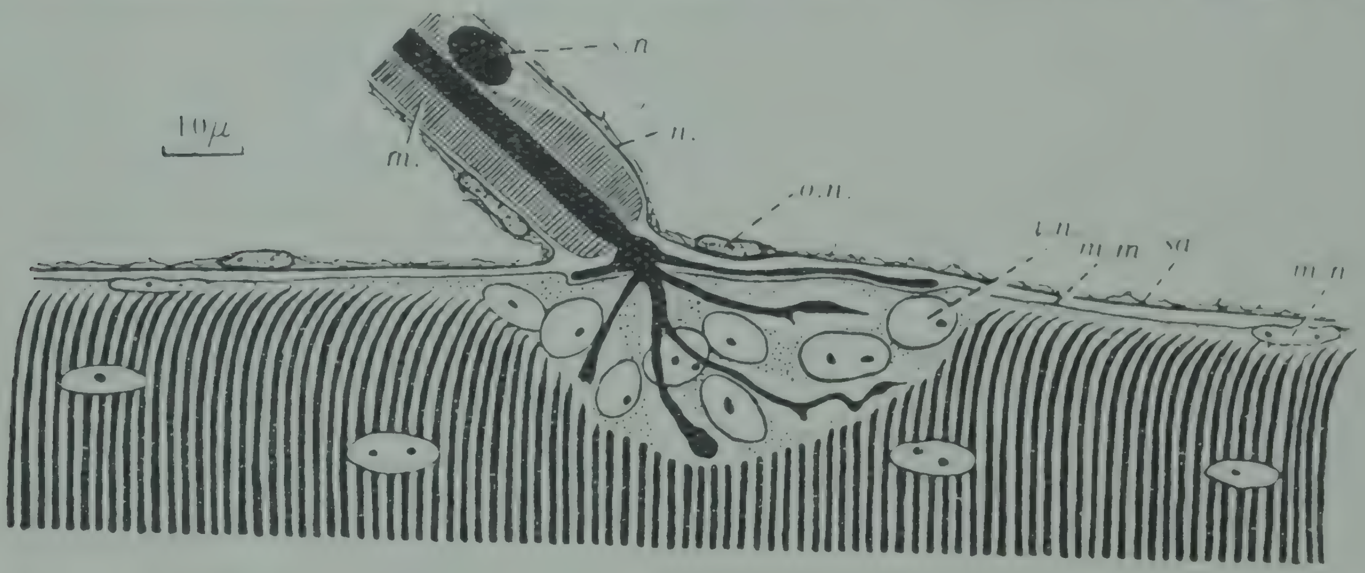


FIG. 32. Semi-diagrammatic view of part of a striated muscle fibre and of its motor end-plate. *m.* myelin sheath; *s.n.* Schwann nucleus; *n.* neurilemma; *o.n.* outer end-plate nucleus; *i.n.* inner end-plate nucleus; *m.m.* muscle (sarcoplasmatic) membrane; *s.a.* sarcolemma; *m.n.* muscle (sub-sarcolemmal) nucleus. (GUTMANN and YOUNG, *J. Anat.*, 1944, **78,** 15.)

The development of this regular cross and longitudinal striation is closely connected with the evolution and specialisation of the muscular function, *i.e.* contraction. In the developing muscle of the embryo or the adult muscles of many invertebrates, this differentiation affects only part of the cell, so that, while one part presents the ordinary granular appearance, the other half is finely and longitudinally striated, due to the development of special contractile fibrillæ. In the striated muscle of vertebrates, and still more in the wonderful wing-muscles of insects which can perform three hundred presumably tetanic contractions in a second, the longitudinal is associated with and subordinated to a transverse striation, due to the regular segmentation of the contractile fibrillæ. The great divergence in the aspect of muscular fibres from different parts of the animal kingdom is largely conditioned by the varying relations, spatial and quantitative, of the sarcoplasm to the fibrils. Thus in the higher vertebrates, two types of voluntary muscular fibre are distinguished, according to the amount of sarcoplasm they contain: one (red fibres) rich in sarcoplasm, more granular in cross section and generally containing a red pigment, myoglobin; the other (pale fibres) poor in sarcoplasm, clear in cross section and containing no myoglobin.

The red fibres are found chiefly in those muscles which have to carry out long continued and powerful contractions. The myoglobin assists in this by acting as a local oxygen store.

In some animals, such as the rabbit, we find muscles consisting almost entirely of one or other of these varieties ; but in most animals the two varieties occur together in one muscle, so that what we have to say about the properties of voluntary muscle really has reference to a muscle containing both red and white fibres. Generally, however, pale muscle predominates in the quick-acting flexor muscles, and red in the tonically acting extensors.

The appearances which are revealed within the fibril by the microscope have been variously interpreted by the eye of faith, or of doubt. The current opinion accepts, at least, the physical reality of the sarcomere and of the striations, though even the permanence of the latter has been questioned.

We will accept the view that the essential contractile element is the chain of sarcomeres forming the muscle fibril (Fig. 33). Each myofibril * is divided by Krause's membrane, often called the Z-disc (the lines in the middle of each light stripe) into sarcomeres. The Z-disc runs right across the fibre from fibril to fibril. Each sarcomere contains a darker substance (A or Q-disc) divided near the centre into two parts by Hensen's (M or Q*h*) disc. At each end of the sarcomere the contents (I or J-disc) are clear and hyaline. Most histologists agree in assigning to the Q-disc of the sarcomere a denser structure than to the two ends. In the resting fibre of the frog the Q stripe occupies 63 per cent. of the height of the sarcomere and the J stripe 37 per cent. When stretched, the Q stripe extends about twice as much as the J stripe.

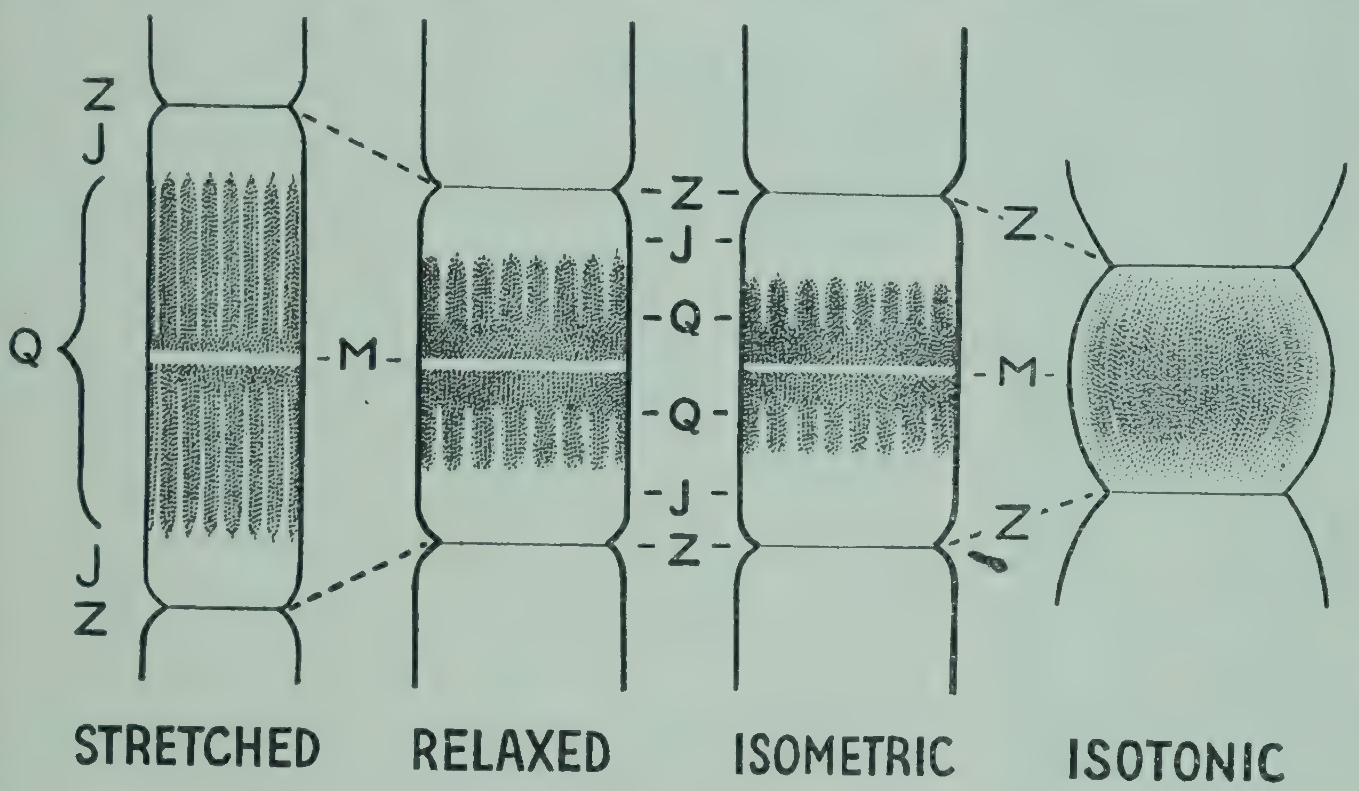


Fig. 33. Diagram of Sarcomeres in various states.

On isometric contraction the sarcomere remains unaltered in length and volume, but there is a shrinkage by 18 per cent. of the dark (Q) stripe and an expansion by 28 per cent. of the light (J) stripe.† Tiegs ‡ considers that the membrane of Krause does not form simple transverse partitions across the fibril, but runs in a spiral form along its whole length as a continuous structure, between the turns of which the sarcomeres are placed (Fig. 34).

When a muscle fibre is examined under the microscope by polarised light, it is seen to be made up of alternate bands of singly and doubly refracting material. The doubly refracting (*anisotropic*) substance which appears bright when viewed with crossed Nicol prisms, corresponds to the dark (or Q) band, and the Z-disc ; the singly refracting (*isotropic*), which now looks dark, to the light (or J) band. Actually the J band is also slightly anisotropic.

Double refraction is due to a substance possessing a regular molecular organisation which is not perfectly symmetrical in all directions. Thus, most crystals show it, but not those which belong to the cubical system, *e.g.*, NaCl. A muscle fibre can be shown by the electron microscope § to be made up of elongated filaments (aggregated molecules) of myosin, each of them birefringent, arranged parallel to one another in the fibril, and running through both light and dark bands in continuity. The transverse bands of cross striation are also visible in these fibrils. It has been found that while solutions of myosin are flowing, or are being stirred, which would orient the elongated molecules parallel to one another, the property of double refraction appears.* This would explain the str birefringence of the Q-discs, but not the much slighter birefringence of the J: it is tr however, that the myosin molecules are more regularly arranged in the Q-discs. It h been stated that the adenosine derivations and lipoids are concentrated in the J ban but most of the cations in the Q band; perhaps these modify the optical properti Glycogen is chiefly found in the sarcoplasm.

---

* Schafer. *Proc. Roy. Soc.*, 1891, **49**, 280 ; *Mt. Monatschr. f. Anat. u. Phys.*, 1891, **8**, 178.

† Buchthal and Lindhard. "Physiology of Striated Muscle Fibres." 1939. (Copenhagen. Munksgaard.)

‡ Tiegs. *Australian J. Exp. Biol. Med. Sci.*, 1924, **1**, 11 ; *Proc. Roy. Soc.*, 1934, **116B**, 38.

§ Hall, Jakus and Schmitt. *Biol. Bull. Woods Hole.*, 1946, **90**, 32.

When a muscle fibre contracts, there is an appare reversal of the situations of the light and dark stripes, owi to the fact that the interstitial sarcoplasm is squeezed o from between the bulging sarcomeres, and accumulates each side of the membranes of Krause. The accumulation sarcoplasm in this situation makes the previously light str appear dark, and the dark striæ by contrast lighter than the were before. That there is no true reversal of the striæ shown by examining the muscle by polarised light, the tw substances, isotropic and anisotropic, retaining their relativ positions, though the double refraction is reduced. The tim course of the change in the double refraction has bee found † to follow a curve very similar to that for tensio development.

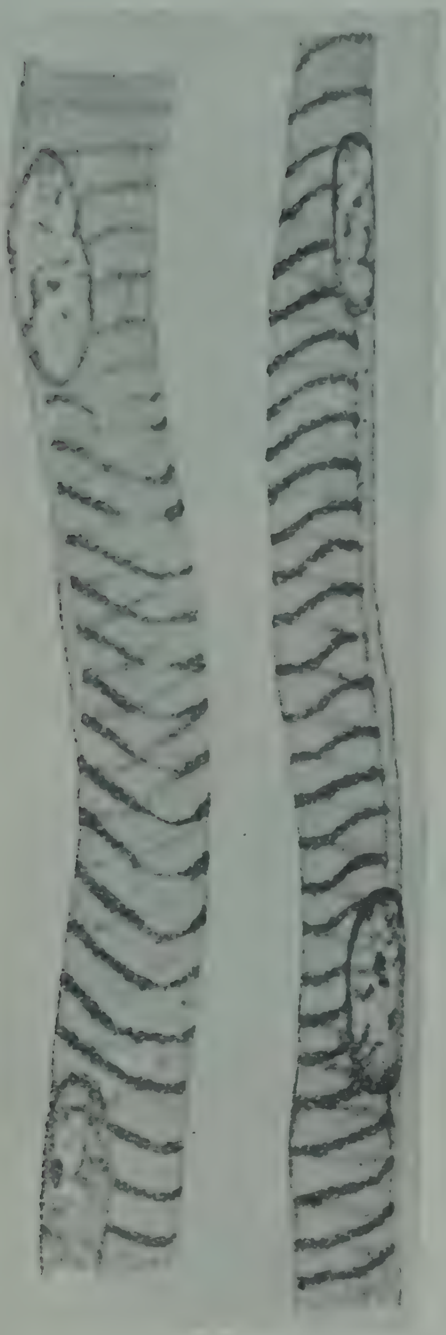


FIG. 34. Two thin muscle fibres of superior oblique muscle of rabbit. The left-hand one shows a double and the right-hand one a single spiral arrangement of striation. (TIEGS. *Proc. Roy. Soc.*, 1934, **116B.**)

Nerve Supply. Every skeletal muscle is connected with the central nervous system by a so-calle motor nerve. The fibres of this are by no mean all motor in function; some, often as many as two fifths of the whole, conveying impressions from th muscle to a centre, some are sympathetic fibres o unknown function, but the remainder are motor fibre conveying nerve impulses from a nerve centre to th muscle. These last—the motor nerve fibres, divide into numerous (sometimes as many as 150) branche when near to the muscle, and each branch terminate in a muscle fibre by means of a special end organ—th motor end plate. Here, the neurilemma of the nerve fibre becomes continuous with the sarcolemma, the medullary sheath ends suddenly, while the axis cylinder ramifies in a plate of nucleated protoplasm, called the *motor end plate*, the protoplasm of which i continuous with a thin sub-sarcolemmal layer of sarcoplasm (Fig. 32). Beneath this layer of sarcoplasm lies the striated contractile substance of the muscle. The ultimate ramifications of the axis cylinder end freely upon the surface of the end-plate protoplasm, but are precluded from physical continuity with this by retaining, and being enclosed by, the thin axolemma of the nerve fibre. This membrane, and the superficial sarcoplastic material, are of considerable theoretical importance in the study of neuromuscular transmission. Each muscle fibre has one end plate. It has been claimed that the end plates are capable of amœboid movements, that when muscles have been exercised they are demonstrably expanded over a larger area,‡ and that the bulk of the sole plate is increased when the muscle is paralysed by curare and diminished when exhausted by strychnine.§

After section of a motor nerve the muscle fibre shrinks and loses its contractility, but the end plates remain intact for many months, and if the nerve regenerates, regain their nerve fibres and may resume function a few days after re-innervation.||

---

* v. MURALT and EDSALL. *J. Biol. Chem.*, 1930, **89**, 315, 351; *Trans. Faraday Soc.*, 1930, **26**, 837.

† v. MURALT. *Pflüger's Arch.*, 1932, **230**, 299.

‡ CAREY. *Amer. J. Path.*, 1942, **18**, 237.

§ CAREY. *Amer. J. Path.*, 1944, **20**, 341.

|| GUTMANN and YOUNG. *J. Anat.*, 1944, **78**, 15.

For Reference

Barer (1948). The Structure of the Muscle Fibre. *Biol. Rev.*, **23**, 159.

Buchthal and Lindhard (1939). The Physiology of Striated Muscle Fibres. (Copenhagen. Munksgaard.)

Hinsey (1934). The Innervation of Skeletal Muscle. *Physiol. Rev.*, **14**, 514.

Jordan (1933). The Structural Changes in Striped Muscle during Contraction. *Physiol. Rev.*, **13**, 301.

Needham, Dorothy (1926). Red and White Muscle. *Physiol. Rev.*, **6**, 1.

Tower (1939). The Reaction of Muscle to Denervation. *Physiol. Rev.*, **19**, 1.

EXCITATION OF MUSCLE AND OTHER TISSUES

In order to avoid subsequent repetitions, we shall now consider the excitation of muscle, and at the same time give a general account of the excitability of other tissues, such as nerve, since many phenomena of excitation are common to all tissues.

A skeletal muscle may be caused to contract in various ways. Normally it contracts only in response to impulses starting in the central nervous system and transmitted down the nerves. But contraction may be artificially excited in various ways in a muscle *in situ* or freshly removed from the body. The agency which is applied to the muscle to evoke a response is called a *stimulus*. Most of the fundamental work on striated muscle has been done with the muscles of cold-blooded animals, mainly the frog, dissected out of the body, often with a length of the motor nerve attached. In such a nerve-muscle preparation we can cause contraction by various forms of stimuli—mechanical, thermal, or electrical—applied to the muscle or to the nerve (direct and indirect stimulation). Thus the muscle responds with a twitch if we pass an induction shock through it or its nerve, or pinch either with a pair of forceps. Or we may cause contraction by the application of strong glycerol or salt solution to the nerve.

When the stimulus is powerful enough to elicit a response it is called an *adequate stimulus* : if not strong enough it is inadequate, but may nevertheless produce certain concealed effects.

Independent Excitability. These experiments do not prove conclusively that muscle itself is excitable. It might be urged that, when we stimulated the muscle, we really only excited the ramifications of the nerve in the muscle, and that these in their turn excited the muscle to contract. This is indeed usually the case,* because nerve fibres are excited by smaller stimuli than are muscle fibres. But that muscle fibres are capable of independent excitability is shown in the classical experiment by Claude Bernard.†

The brain of a frog is destroyed, and the sciatic nerves on each side exposed. A ligature is then passed round the right thigh underneath the nerve, and tied tightly so as effectually to close all the blood vessels supplying the limbs, without interfering with the blood supply, to the nerve. A little curare solution is then injected into the dorsal lymph sac. After a quarter of an hour it is found that the strongest stimuli may be applied to the left sciatic nerve without causing any contraction of the muscles it supplies. On the right side, stimulation of the nerve is as efficacious as before. Both gastrocnemii respond readily to direct stimulation, showing that the muscles are not affected by the drug. Since both sciatic nerves have been exposed to the influence of the curare, the difference on the two sides cannot be due to any effect on them of the curare. We have also excluded the muscles themselves. So we must conclude that the curare paralyses the muscles by preventing the excitatory process from passing over from the nerve to the muscle, and that muscle can be excited by direct stimulation.

Direct proof of independent excitability of muscle is provided by the fact that individual muscle fibres may be seen under the microscope to contract when stimulated by microelectrodes at points remote from end plates.‡

ELECTRICAL EXCITATION

It is characteristic of excitability in general that a stimulus of very small energy content may evoke a response in which very much larger amounts of energy are released.

---

* Asmussen. *Strand. Arch. f. Physiol.*, 1933, **65**, 261.

† Claude Bernard. Substances toxiques, p. 344 (Paris, 1857). *Compt. rend.* 1856, **43**, 825.

‡ Gelfan. *Amer. J. Physiol.*, 1930, **93**, 1.

Of all the different stimuli capable of exciting muscular contraction, th electrical is the most satisfactory, because the phenomenon of excitation is itse of an electrical nature. It is easy to graduate accurately the intensity an duration of the stimulus, and the stimulus may be applied many times to an point on the muscle or nerve without injury. The four commonest forms c electrical stimuli employed are (1) the make and break of a constant current such as that from an accumulator, (2) the inducted currents of high intensit and short duration obtained from an induction coil, (3) condenser discharge (4) alternating current.* As the student will become familiar with thes in his practical work, no description of the common forms of apparatus neede for producing these stimuli will be necessary here.†

Resistance and Impedance. When a constant current is passed through a tissue, the *resistance*, in accordance with Ohm's law, is,

$$R = E/C,$$

where E is the e.m.f. and C the current intensity. Owing to polarization (p. 116) the current strength rapidly falls off to a low value, because the permeability of the membranes in the tissue to the ions is restricted.

If an alternating current is used less resistance is offered because of the absence of polarisation. The resistance to an alternating current is called *impedance*. With an alternating current $I = I_0 \sin \omega t$, where $\omega = 2\pi \times \nu$ (ν = frequency and I_0 = amplitude) the impedance Z includes resistance R and capacitance C. If resistance and capacitance are in parallel, then ‡

$$Z = \sqrt{\frac{R^2 + \omega^2C^2R^4}{(1 + \omega^2C^2R^2)^2}}$$

Galvanic Stimulation. If a weak *constant current* be passed through a muscle or any part of its nerve, at the make of the current the muscle gives a single sharp contraction—a *muscle twitch*, in which, according to the current strength, and other conditions, few or many of its muscle fibres may be involved. During the passage of the current no effect is apparently produced, though on careful observation we may with strong currents see that there is a state of continued contracture limited to the immediate neighbourhood of the cathode, which lasts as long as the current is passing through the muscle, and is not propagated to the rest of the muscle. If the current be now broken, the muscle may remain quiescent, but if the current was above a certain strength, the muscle responds to the break of the current with another single twitch. With a current of moderate strength we may get a contraction both at make and break of the current, *the make stimulus being more effective than the break stimulus.*

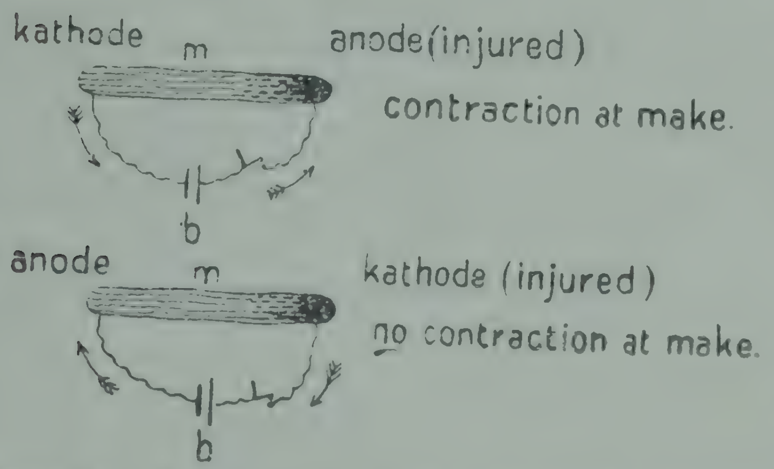


Fig. 35. Diagram to show the Effect of Local Injury on the Excitability of a Muscle. *b*, battery; *m*, muscle. The arrows indicate the direction of the current.

Besides this difference in intensity at make and break of the constant current, there is a difference in the point from which excitation starts. *A make contraction starts from the cathode, a break contraction from the anode.* This is called *polar stimulation* and is well shown by the following experiment.

If a muscle be killed at any spot, its excitability at this spot will be destroyed. Hence if we stimulate a muscle at the dead spot, no contraction will ensue. This fact may be used to demonstrate the facts of polar excitation. A muscle with parallel fibres such as the sartorius, is killed at one end, and a constant current passed, first from the dead to the uninjured end, and then in

---

[*] Katz. *J. Physiol.*, 1939, **96**, 202.

[†] *See* D. T. Harris. "Experimental Physiology," 1941. (Churchill.) For a precision stimulating device (thyratron tube), *see* Schmitt and Schmitt. *Science* N. S. 1932, **76**, 328.

[‡] Hubbard and Rothschild. *Proc. Roy. Soc.*, 1939, **127B**, 510.

the reverse direction (Fig. 35). It is found in the former case, when the anode is on the dead part (which is therefore inexcitable), that break of the current is ineffective, and in the latter, when the cathode is on the dead surface, that the make stimulus is ineffective, showing that the part excited corresponds to the cathode at make and to the anode at break.

As will be seen later, the *duration* of the flow of current also affects the result when short periods of flow are concerned. With a current of very short duration no excitation is produced at break. Every induction shock is such a current of very short duration produced in the secondary coil by make or break of the primary circuit, and can be therefore regarded as a make stimulus, and when such a shock is led through a muscle the contraction in each case will start at the cathode, *i.e.* the point at which the induction shock leaves the muscle.

The general laws for stimulation of nerve fibres are similar to those for stimulation of muscle fibres directly.

Faradic Stimulation. Using the *induced current* as a stimulus, it is found that the contraction on break of the primary current is often stronger than that on make. It must not be imagined, however, that there is any contradiction between this and the fact that the make of a constant current is a stronger stimulus than the break. It is due to the fact that the break-induced current is, for physical reasons, a stronger one than the make current. Both are currents of very short duration, so short that only the rising phase of each is effective. It has been shown that, when we use currents of such short duration, the fall is ineffective ; so in both cases, whether we make or break the current in the primary circuit, we are dealing with a *make* stimulus in the muscle.

In using either of these modes of stimulation we find that there is a certain intensity which the stimulating current must possess in order that any effect shall be produced. Any strength of stimulus below this is known as a *subliminal* stimulus. A *liminal* stimulus (sometimes known as *minimal* or *threshold* stimulus) is the weakest stimulus that will produce any result, *i.e.* in muscle a contraction. With galvanic excitation, this is also called the *rheobasic strength.* A *maximal* stimulus is one that produces the strongest contraction of which a muscle is capable as a result of a single stimulus. A *submaximal* stimulus is any strength of stimulus between these two extremes.

Time Factor in Excitability. Examination of the properties of different living tissues demonstrates that their excitabilities vary widely. Some tissues require on galvanic stimulation a much stronger current than others, or, on faradic stimulation, a closer coupling of the secondary to the primary coil. Again, some respond to galvanic but not to faradic excitation, or *vice versa.* It is clear that comparison of excitabilities in different tissues is unsatisfactory, if at all possible, on such indications as these. In view of the fact that there is no fundamental qualitative difference between galvanic and faradic currents, such a statement as " excitable to faradic, inexcitable to galvanic electricity," is meaningless.

Dealing with galvanic stimulation as more readily analysable than faradic, it is found that the following must be taken into account in studying excitability :—

(*a*) *The current density* or strength of current (c) or else the *voltage* (V)—to which (c) is proportional—required to stimulate.

(*b*) *The duration of flow* of the current. It has been found that a current that will just excite when of a particular duration, fails to do so when its duration is reduced below a certain level ; also, for a given current, further continuation of the current, beyond a particular duration, differing for each tissue, produces no further effect.

An elementary illustration, given by Lapicque, will serve to show the importance of the time factor in excitation. A liquid potentiometer is used consisting of a long narrow trough filled with zinc sulphate solution. Into it dip two plates of amalgamated zinc A and B (Fig. 36), placed about one metre apart and connected with the source of current ; two other similar plates also dip into the gutter; O is stationary and near to A, while C can be dipped in and out of the solution, thus making and breaking contact. The electrodes EE are placed in contact with a strip of plain muscle (frog's stomach). The strength of current flowing through the tissue is proportional to the distance of C from O. If C be plunged into the groove at intervals of a few seconds, farther and farther from O at each immersion, a point, the *threshold* will be reached, at which each prolonged immersion is followed by contraction of the plain muscle. The distance OC at which this occurs (say 10 cm.) is noted. Now C is withdrawn and dipped as quickly as possible in and out of the groove at the same point, giving stimuli lasting only about one-tenth of a second; no contraction will result. If the distance OC is made greater, however (say 70 cm.), so as to increase the potential tapped off, even this brief duration will stimulate.

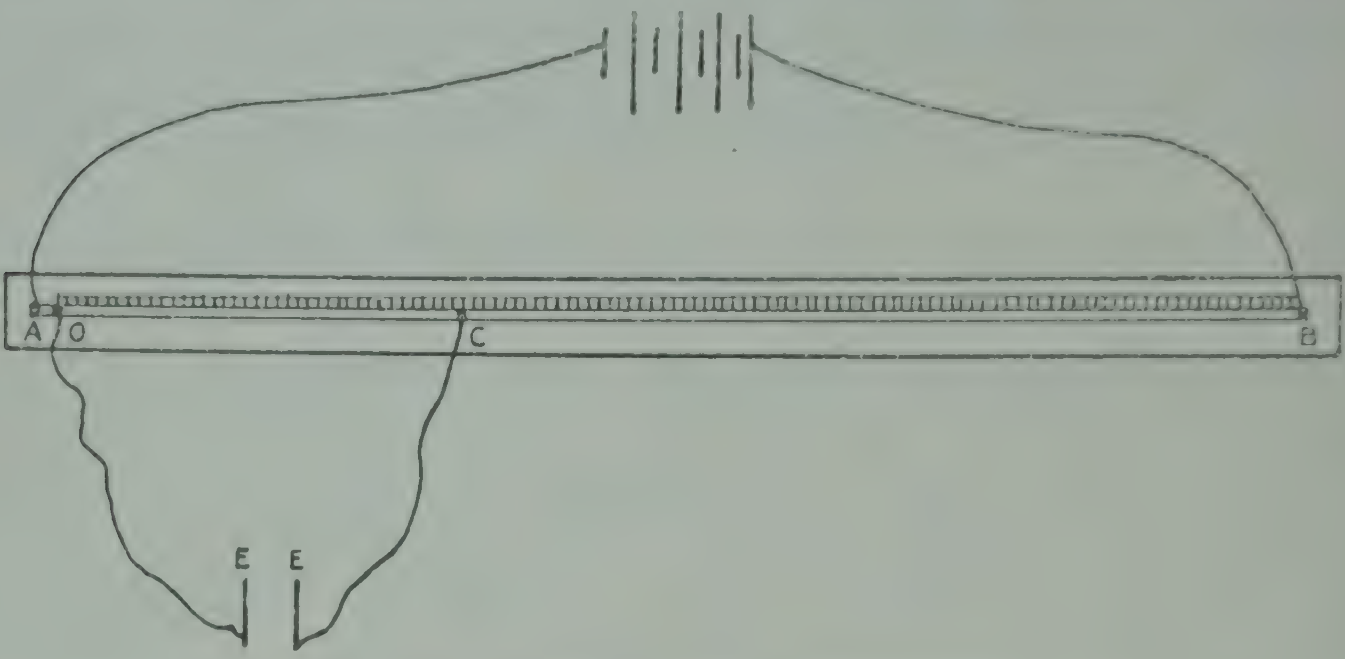


FIG. 36. Simple Liquid Potentiometer.

(*c*) The *rate of change* of the current strength. In (*a*) and (*b*) it has been assumed that the current was established and stopped instantaneously ("rectilinear" or "square" waves) as in Fig. 37 (A), whereas under certain circumstances a current might fluctuate by a slope, *e.g.* as shown in Fig. 37 (B), and these fluctuations would need to be taken into account. If, instead of plain muscle, the nerve of a frog's sciatic-gastrocnemius preparation be placed in the electrodes EE, and the cursor C of Fig. 36 be *quickly* moved along the groove, say, from O to 20 cm., so as to cause a rapidly increasing current to flow through the nerve, the latter will be stimulated as shown by contraction of the muscle. Now return the cursor to O and move it away *very slowly*; it will be found that although it be moved far beyond the previous distance (say to 50 cm.), no contraction will follow—the slow slope of current-increase is ineffective, whereas the steep gradient stimulated the nerve.

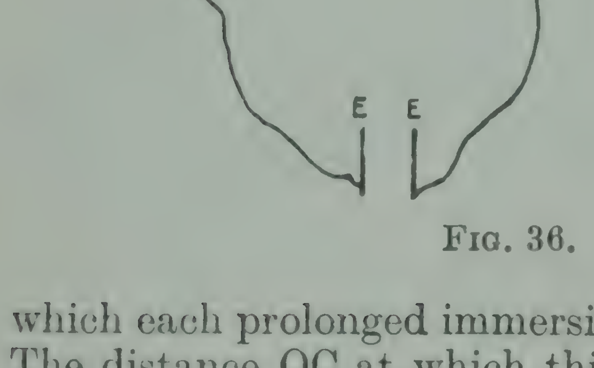


FIG. 37.

Du Bois Reymond expressed this fact by stating that excitation depends on the rate of variation of current density from moment to moment; the more rapid the change the greater the excitatory effect.

This statement is now known to be only partially true. We should now regard it as more correct to say that the current excites during the first moments of its passage, whether variable or constant, but that a certain duration of flow is essential.

To demonstrate the importance of duration in the stimulation of such highly excitable tissues as nerve or striated muscle we require much more refined apparatus than that described above for use with the relatively inexcitable plain muscle. If we can get currents of sufficiently short duration, we can demonstrate that with *any* tissue a liminal current which will excite at a longer duration fails to do so when it is abbreviated beyond a certain limiting value. This limit for the frog's stomach is of the order of tenths of a second, while for the sciatic nerve it is of the order of thousandths of a second.

These simple considerations serve to illustrate the importance of the time factor in excitation, and to show that "slow" and "rapid" are purely relative terms when different tissues are under consideration.

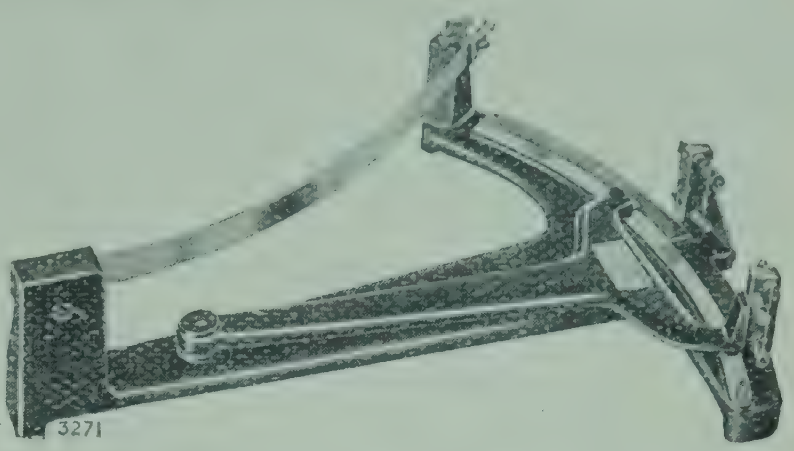


FIG. 38. Keith Lucas's Spring Rheotome.

Currents of sufficiently short duration can be obtained by the use of a mechanical device or rheotome, such as Lucas's pendulum or spring (Fig. 38), by which two keys are opened at known intervals. The first key is a short circuiting key, the opening of which starts the current, and the second one is in series, and its opening stops the current.

Another method is by the use of condenser discharges. The time taken for a condenser to discharge can be calculated if its capacity and the resistance of the circuit are known, the time being proportional to the product RC. The current strength is, of course, not constant during the discharge, but it is much easier to obtain currents of very short duration by the adjustment of condenser capacities than by any of the rheotomes for use with constant currents, and the drawbacks attendant on the use of discharges of rapidly declining strength have been overcome, since it has been shown that a discharge from a condenser charged to a given voltage is equivalent to a rectilinear constant current of duration = 0·37 RC seconds, at the same voltage (R = resistance of circuit in ohms ; C = capacity of condenser in farads).

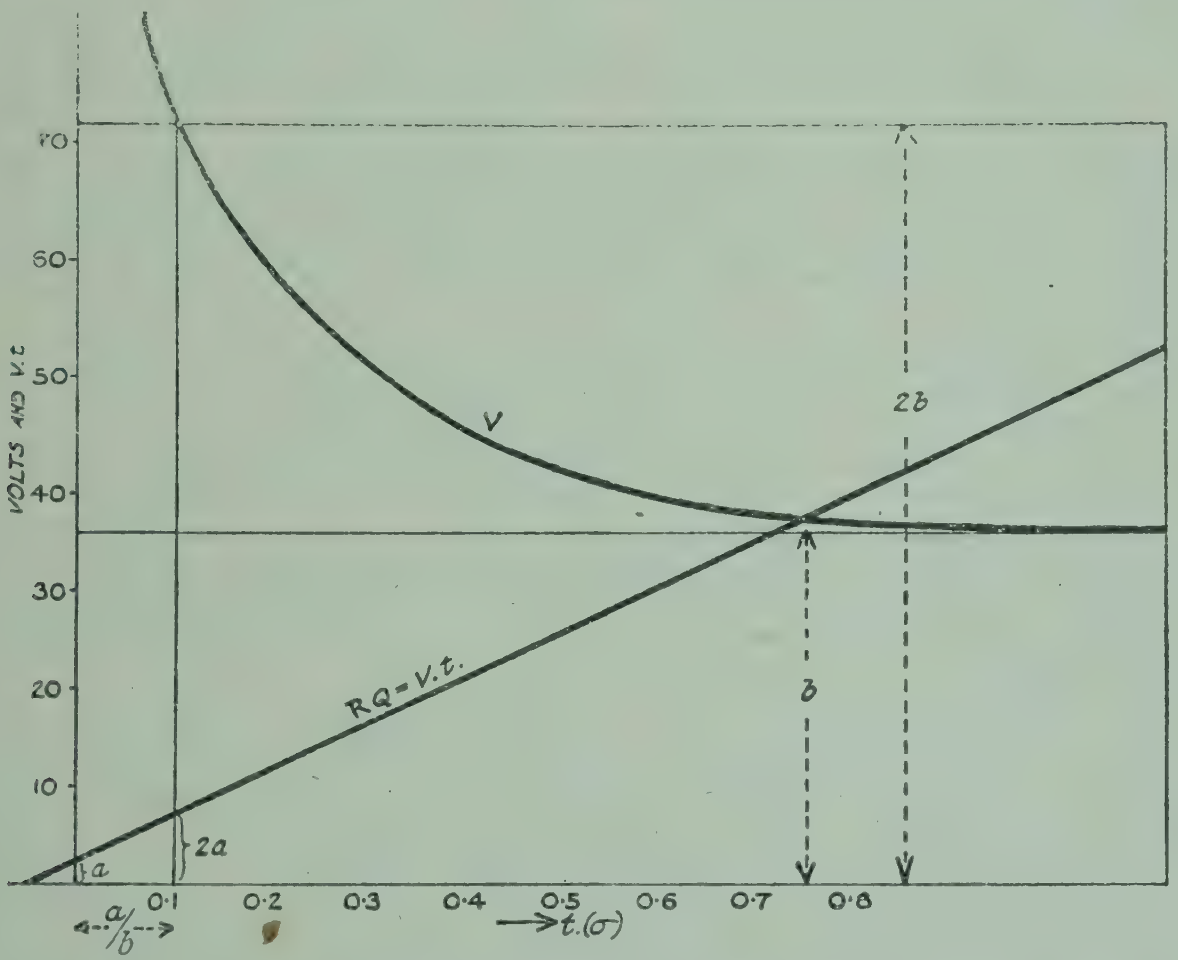


FIG. 39. To illustrate Weiss's Laws and their Relation to Chronaxie. V = curve of voltages. RQ = curve of quantities. $a = 3{\cdot}5$. $b = 36$ volts. $\frac{b}{a} = \bar{c}\ 0{\cdot}1$ msec. = chronaxie. Abscissæ = time in 0·1 msec. (σ). Ordinates figured = volts. Ordinates for RQ = V*t* are on same scale. (Modified from BOURGUIGNON. *La Chronaxie.*)

When the smallest voltages which will excite a tissue are plotted against the durations of the current, controlled in either of these ways, a curve of the rectangular hyperbola type (Fig. 39) is obtained, one of the asymptotes of which is a line at the level of the threshold voltage (*b*) for currents of infinite duration. Thus in Fig. 39, a voltage lower than *b* (= 36 v.)

will not stimulate that particular tissue under the given conditions of experiment ; a curren of 40 v. will stimulate if it lasts for not less than 0·55 msec., one of 50 v. if it lasts for a least 0·35 msec., etc.

For a different tissue, and with other conditions of experiment, quite different voltage and times would probably be obtained, but the form of the curve would always be the same, though the scales would differ. Such curves are called *strength-duration curves*. An important condition in the preparation of these is that all the conditions of experiment, except strength and duration of the exciting current, must remain the same throughout any given curve.

If the quantities of electricity (proportional to Vt) are plotted against times, a straight line is obtained.

These relations for constant currents are expressed by Weiss's laws :—

(1) For intensity,

$$= \frac{V}{R} = \frac{a}{i} + b.$$

(2) For quantity,

$$Q = it = a + bt,$$

where a and b are two constants.

a is the quantity of electricity where the graph cuts the ordinate.

b is the threshold strength at infinitely long duration. Almost identical relations are obtained by the use of condensers (Hoorweg's laws).

In a strength-duration curve, current strengths are plotted against the *excitation-times*, *i.e.* against the shortest times in which each strength can excite. Evidently, if other things were equal, the more excitable a tissue the shorter the time needed to rouse it to respond.

Lapicque calls the liminal value for indefinite duration (galvanic threshold) the *rheobase*, and estimates the excitability of a tissue by what he has called the *chronaxie*. This is one specially defined excitation-time.

The chronaxie may be defined as the shortest duration of a current of twice the rheobasic strength which will stimulate. Thus in Fig. 39, when the voltage is raised to 2b, the shortest duration which will excite is 0·1 msec., and this duration is the chronaxie. A less excitable tissue has a longer, and a more excitable one a shorter chronaxie.

In practice, in order to determine the chronaxie, all that is necessary is to find the galvanic threshold or rheobase, using an ordinary key, then double the voltage so employed, and finally, using either condensers or a suitable rheotome, find the shortest duration of this which will excite.

Provided the conditions of experiment are constant, the chronaxie of a tissue is a definite measure of its excitability, since in this time a current of twice the rheobasic strength is able to produce those physico-chemical alterations which determine excitation, and the more excitable the tissue is, the more quickly will that current produce those changes. It can be shown that what happens in an excitable tissue in, say, 0·3 msec., may require, say, eight seconds for a very inexcitable one.

Whether it is feasible to obtain constant conditions from one experiment to another, or even throughout a single experiment, is, however, a matter of opinion, which has been contested with some heat. Variation in the size or nature of the stimulating electrodes is especially liable to produce widely different results.*

The following table shows some approximate chronaxies for excitable tissues of very varying types :—

Tissue.	Chronaxie msec.	Approximate Duration of Contraction of Muscles.
		Seconds.
Frog's sciatic and gastrocnemius (esculenta) .	0·3	0·1
,, ,, ,, (temporaria)	0·7	0·1
Toad's sciatic gastrocnemius . . .	1·3	0·16
Frog's ventricle	3·5	1·0–2·0
Tortoise heart (testudo græca) . . .	8·2	2–3
Claw muscle of crab	30	5
Frog stomach	30–100	15–20
Pigment cells, skin of frog	1,000–15,000	—

It is evident from consideration of the figures for different chronaxies that there is a relation between chronaxie and speed of action of tissue. The more excitable tissue has, as a rule, a shorter chronaxie, shorter period of action, shorter refractory period, and a quicker rate of propagation than a less excitable one.

It is because of its very short duration (0·1 msec.) that the induced current is unable to effect the excitation of sluggish tissues such as some forms of plain muscle. When the duration is made extremely short, even very high voltages fail to excite any tissues. Hence high frequency alternating currents of enormous voltage can be passed through the body without any physiological effect other than that due to the heating effects, as in the practice of diathermy.

---

* For a criticism of the experimental facts regarding chronaxie, see RUSHTON. *Journ. Physiol.*, 1932, **74**, 424 ; **75**, 445.

According to Lapicque, the same chronaxie is found for a muscle stimulated directly or *via* its nerve. The classical view, which has been upheld by Rushton, was that the muscle is much less excitable than its motor nerve, and that when the " muscle " is stimulated it is really the nerves within it which are excited and these then stimulate the muscle fibres. In support of this opinion he adduces the α and γ curves of the nerve-muscle complex which were discovered by Keith Lucas. This observer found that the strength-duration curve of a muscle (with its contained nerve branches) was composite, as shown for curve B (Fig. 40), and supposed this to be due to the presence of a more excitable (or γ) tissue, which was probably nerve, and of a less excitable one (α), probably muscle. Lapicque, however, obtained a uniform curve resembling C, of properties akin to Lucas's α-curve. Rushton showed that in a parallel-fibred muscle either α- or γ-, or composite curves could be obtained, according to the angle at which the stimulating current passed relatively to the direction of the muscle and nerve fibres. This is because muscle and nerve fibres are most excitable when stimulated by currents parallel to their fibres, and least excitable to currents running at 90° to their fibre direction (more precisely the rheobase varies inversely as the cosine of the angle between current direction and fibre direction). Whether a given current tends to excite the muscle fibres or the intra-muscular nerve fibres will therefore depend not only on its strength and duration, but also on its relation to the direction of the muscle and nerve fibres. In many muscles a majority of the small nerves run roughly at right angles to the muscle fibres (Fig. 41). Hence, if as in Fig. 40 C, we stimulate with a current parallel to the muscle fibres we get the muscle or α curve, if at 90° the γ or nerve-excitation curve (Fig. 40 A), and if at 45° a composite curve (B).

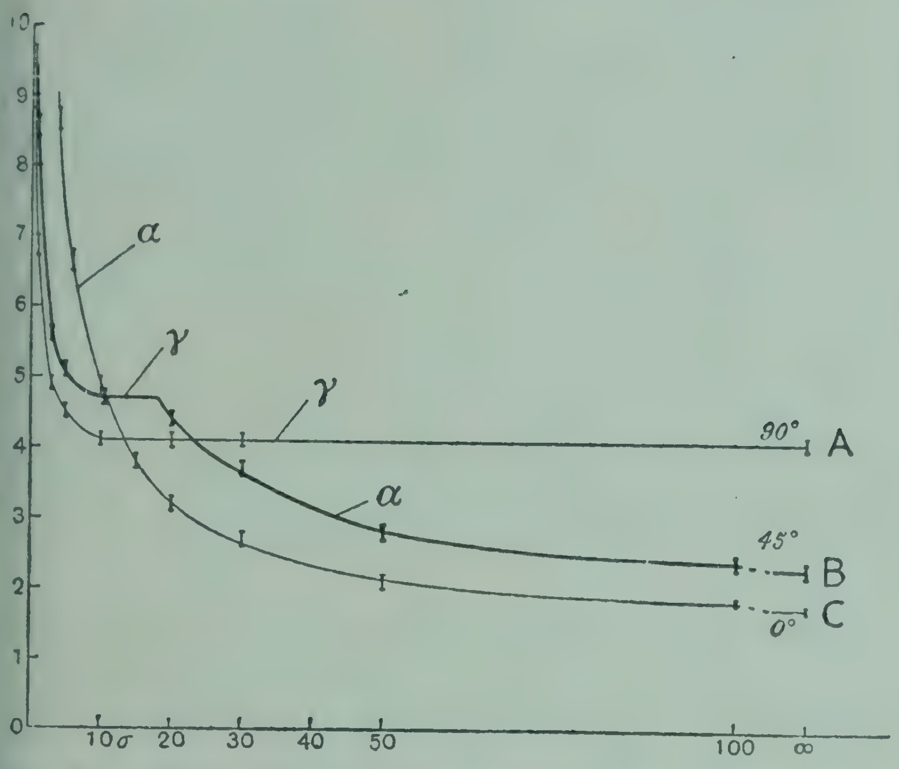


FIG. 40. Strength-duration curves from sterno-cutaneous muscle of frog stimulated at 90°, 45° and 0° to direction of muscle fibres. (RUSHTON, *J. Physiol.*, 1930, **70**, 334.)

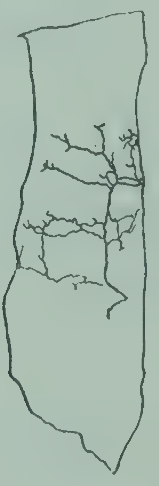


FIG. 41. Sterno-cutaneous muscle of frog showing direction of nerve fibres. (RUSHTON, *J. Physiol.*, 1932, **75**, 174.)

If the intra-muscular nerves are dissected away, or if curare is given, the curve becomes a smooth α-curve, characteristic of muscle, under all circumstances.

Lapicque's explanation is entirely different. He believed that unless the nerve and muscle have *approximately* the same chronaxie the excitatory state will not be transmitted (law of isochronism): this failure occurs when the chronaxie of one tissue is more than twice that of the other. Curare, he claimed, increases the chronaxie of muscle but not that of nerve; strychnine on the other hand reduces that of nerve but leaves the muscle unaltered: in both cases conductivity from nerve to muscle fails. He claimed that the conditions of Rushton's experiments are not such as would be expected to give true chronaxies.

LATENT ADDITION. Closely associated with the excitability of the tissues are the phenomena of latent addition of excitation, or " *summation of stimuli* " and of " *refractory period.*" If two *subliminal* stimuli are sent in within a sufficiently short interval of time, heir effect is summated so that two stimuli, each of which would be ineffective, may ogether produce an excitation. For summation of two stimuli to take place, the second timulus must occur at a time before the condition excited by the first stimulus has died way. The maximum time interval at which summation can take place will therefore vary rom tissue to tissue, and will bear a relation to the chronaxie. This is evident if we compare the maximum summation intervals for different tissues with their chronaxies. Thus for frog's nerve, with a chronaxie of 0·3 msec., the summation interval is 0·5 msec., while for frog's heart, with a chronaxie of about 3·5 msec., the maximal summation interval is 8·0 msec.

For Reference

Davis and Forbes (1936). Chronaxie. *Physiol. Rev.*, **16**, 407.

Fredericq (1928). Chronaxie. *Physiol. Rev.*, **8**, 501.

Lapicque (1926). L'excitabilité en fonction du temps. (Paris. Press Universitaires de France.)

Lapicque (1934). *Journ. Physiol.*, **81**, 113.

McIntyre (1947). Some Physiological Effects of Curare and their Application to Clinica Medicine. *Physiol. Rev.*, **27**, 464.

Rushton (1933). *Journ. Physiol.*, **77**, 337.

ELECTRICAL CHANGES IN MUSCLE

The fact that electrical currents can be obtained from living tissues was proved by Galvani * by his demonstration of " contraction without metals," as described below. It is easy with present-day apparatus to show that all living tissues when active exhibit differences of electrical potential, and these phenomena are of considerable practical and theoretical importance. Much of the description given here for muscle applies equally well to other tissues.

Polarisation. If a current from a battery be passed between two plates of platinum immersed in acidulated water or salt solution, *electrolysis* of the water takes place, bubbles of oxygen appearing on the anode, and bubbles of hydrogen on the cathode. If now we remove the battery, and connect the two electrodes by wires with a galvanometer, a current passes through the galvanometer and water in the reverse direction to the previous battery current. This current is called the polarisation current. The vessel in which the electrodes are immersed has in fact become a storage cell, the platinum covered with oxygen bubbles being the positive pole, and that covered with hydrogen bubbles the negative pole. Similar polarisation takes place when we pass currents through the tissues of the body by means of metallic electrodes.

We have already learnt in Chapter VIII that the electrical phenomena of living tissues are due to the existence of polarised membranes at the surfaces of the cells ; *i.e.* of membranes having a positive charge on one surface, usually the outer one, and a negative charge on the other surface. It is therefore essential in studying such phenomena to avoid other forms of polarisation, which would mask or modify them.

Hence, before we can pass currents into, or lead off currents from, living tissues, it is necessary to have some form of electrodes in which this polarisation will not occur. With one exception, polarisation always results when a current passes from an ordinary metallic conductor to a saline solution, but not when it passes from one saline solution to another. This exception is when the metal dips into a solution of its own ions, and this fact is taken advantage of in the construction of *non-polarisable electrodes.* In these a metal dips into a solution of its salt ; any tendency for polarisation is at once counteracted at the cathode by deposition of the metal on to the electrode from the solution, or at the anode by solution of the metal of the electrode to neutralise the negative charges carried thence by anions (Fig. 42). Some types of these electrodes are shown diagrammatically in Fig. 43.

In such electrodes the conduction of the current through the nerve or muscle to the metallic part of the circuit may be presented as shown in Fig. 42.

If a muscle such as the sartorius be removed from the body, and two non-polarisable electrodes connected with a delicate galvanometer be applied to two points of its surface, there will be a deflection of the galvanometer, showing the presence of a current which flows in the external circuit from uninjured to injured surfaces. This current of resting muscle is due to the leading off from the cell interiors at the injured places and the cell surfaces at uninjured ones, as explained in Chapter VIII. Hermann described the currents of rest thus : " In partially injured muscles every point of the injured part is negative towards the points of the uninjured surface." These phenomena are explained in Chapter VIII, Figs. 27—30. When the injured fibres are quite dead, this current of rest or " demarcation current " disappears. The potential is due to the electrical charge on the two sides of the plasma membrane, negative inside and positive outside. If the sartorius of the frog be cut out and immersed for

---

* The publication of 1794, the famous " *Dell' arco conduttore,*" was anonymous. v. Fulton and Cushing. *Annals of Science*, 1936, I. 239.

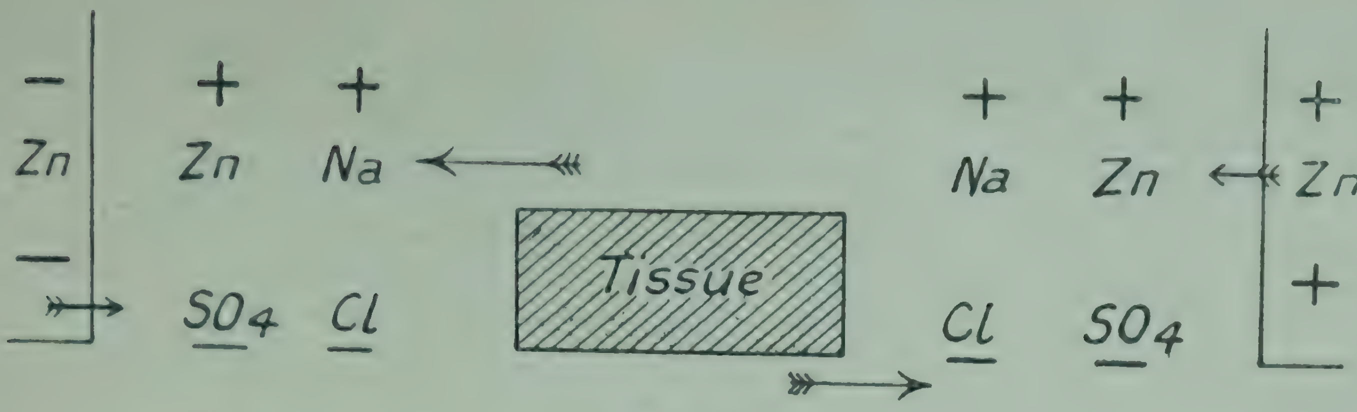


FIG 42.

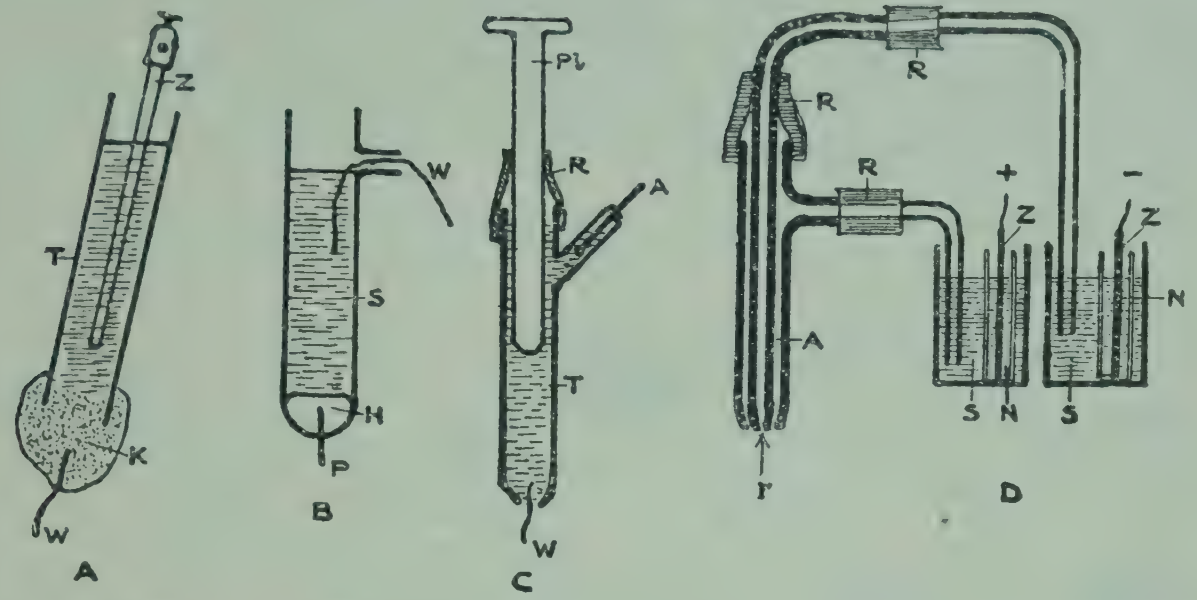


FIG. 43. Types of Non-polarisable Electrodes.

(A) Noyons' type of non-polarisable electrode. Z = amalgamated zinc rod imbedded in 10 per cent. gelatine containing $ZnSO_4$ and contained in the glass tube, T. K = plug of kaolin made with isotonic NaCl solution. W = thread wick soaked in isotonic NaCl leading to tissue.

(B) Ostwald type of N.P.E. P = platinum wire making contact with mercury H. S = isotonic NaCl solution containing solid calomel. W = saline wick lead to tissue.

(C) Lapicque's AgCl electrode. A = silver electrode coated with AgCl. Pl = plunger fitting rubber sleeve R, for filling electrode with NaCl solution. W = saline wick lead.

(D) Pratt's pore electrode. r = the fine pore which is the stimulating point. The tube is filled with saline solution and connected with the saline solution in the vessel S, in which is a porous N.P.E. cell filled with $ZnSO_4$ solution and having a Zn rod immersed in it. The outer tube, A, is the other electrode and is similarly connected with the other N.P.E. cell. R = rubber connections.

twenty-four hours in Ringer's solution, all the injured fibres die, and the uninjured fibre surfaces are then found to be equipotential.

If one end of the muscle be now injured by cutting or burning, a galvanometer will show the presence of a potential between injured and uninjured tissue. The existence of this current may be demonstrated as in Galvani's classical experiment, without using a galvanometer. If the nerve of a sensitive muscle-nerve preparation *a* (Fig. 44) be allowed to fall on an excised muscle *b*, so that one point of the nerve is in contact with the cut end and one with the surface of the second muscle *b*, the muscle *a* will contract each time the nerve touches *b* so as to complete the circuit. If the nerve of *b* is tetanised, *a* as well as *b* enters into a tetanic contraction, for reasons that will be clear later.

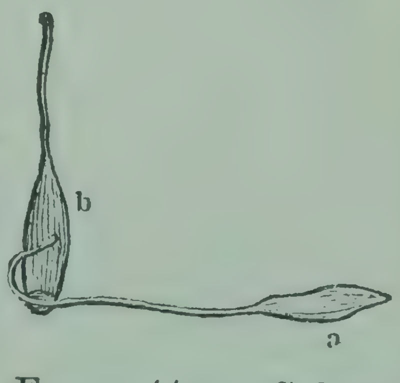


FIG. 44. Galvani's "Contraction without metals."

Although resting uninjured muscle shows no current, a definite electrical change occurs in it when it contracts. As we saw in Chapter VIII these changes are due to depolarisation (or even reversed polarisation) of the membranes at the active places. To show this change, we may lead off two points, one on a cut end and one on the surface of the muscle of a muscle-nerve preparation, to a galvanometer. We shall first obtain a deflection of the mirror due to the "*current of rest*," the cut surface being negative to the resting uninjured surface. If now the nerve be stimulated in any way, so as to throw the muscle into a tetanus, the galvanometer mirror swings back towards the zero, showing that the current which was present before is diminished, *i.e.* the uninjured surface on excitation also becomes negative, so that little or no current now flows from it to the injured part. When the excitation of the nerve is discontinued, the galvanometer indicates once more the original current of rest. This diminution of the current of rest during activity used to be spoken of as the "negative variation." It is now usually termed an "action potential."

In order to obtain a graphic record showing the time-relations of the electrical changes accompanying a single muscle twitch, it is necessary to employ some instrument which can react much more rapidly than the ordinary galvanometer. For this purpose we may employ either the capillary electrometer, the string galvanometer or a suitable oscillograph.

The Capillary Electrometer consists of a glass tube drawn out to a fine capillary, and filled with mercury. The point dips into dilute sulphuric acid. The meniscus of the mercury in the capillary is observed under the microscope. If now the capillary and acid be connected with two points, any difference in the potential of these two points causes a movement of the meniscus in the direction in which the current is flowing, and the extent of the excursion varies with the difference of potential. It is little used to-day.

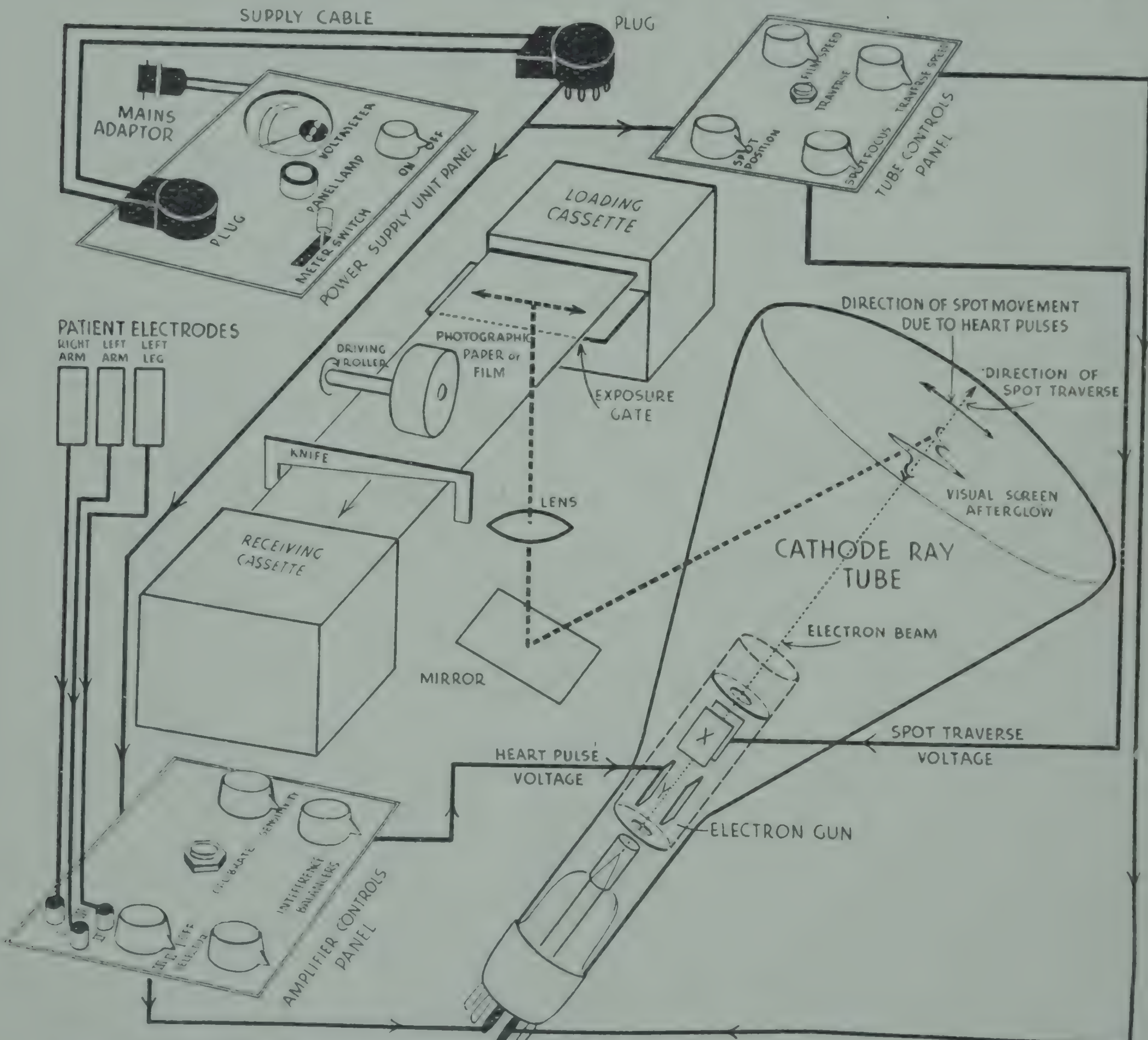

SUPPLY CABLE
PLUG
MAINS ADAPTOR
VOLTMETER
ON
OFF
PANEL LAMP
PLUG
METER SWITCH
POWER SUPPLY UNIT PANEL
FILM SPEED
TRAVERSE
TRAVERSE SPEED
SPOT POSITION
SPOT FOCUS
TUBE CONTROLS PANEL
LOADING CASSETTE
PATIENT ELECTRODES
RIGHT ARM
LEFT ARM
LEFT LEG
DRIVING ROLLER
PHOTOGRAPHIC PAPER or FILM
EXPOSURE GATE
DIRECTION OF SPOT MOVEMENT DUE TO HEART PULSES
DIRECTION OF SPOT TRAVERSE
KNIFE
LENS
VISUAL SCREEN AFTERGLOW
RECEIVING CASSETTE
CATHODE RAY TUBE
ELECTRON BEAM
MIRROR
X
SPOT TRAVERSE VOLTAGE
HEART PULSE VOLTAGE
ELECTRON GUN
SENSITIVITY
CALIBRATE
INTERFERENCE BALANCERS
AMPLIFIER CONTROLS PANEL
SUPPLY VOLTAGES

FIG. 45. Diagram of arrangement of cathode ray oscillograph with photographic apparatus, for electrocardiography. For other purposes a similar arrangement can be used, lead-off being to any two of the "patient electrodes" shown in the figure. (HARRIS.)

In the *String Galvanometer of Einthoven* a very delicate thread, of silvered quartz or of platinum, is stretched between the poles of a strong magnet, which are pierced by holes so that the thread may be illumined by an electric light from one side, and from the other a magnified image of the thread may be thrown upon a screen. Whenever a current passes through the thread it moves laterally, and the lateral movement may be photographed on a moving photographic plate. It will detect very small currents and will respond reasonably accurately to very rapid changes in potential. The records of rapid fluctuations are more nearly correct than those obtained by the capillary electrometer, but such errors as are present are less amenable to correction.

All modern workers prefer to use the more delicate and more rapidly moving *oscillograph*, such as the *cathode ray oscillograph*, the potentials being amplified by valves, and the movements of the beam being photographed (Fig. 45). This instrument is an electrometer (measurer of difference of potential), and not a galvanometer (current measurer) so that no current passes the surface. Hence the use of non-polarisable electrodes is not essential.

Diphasic and Monophasic Potential Waves. If an uninjured regular muscle (Fig. 46) such as the sartorius be kept in moist air and stimulated with a single induction shock at one end, *x*, or indirectly, and two points, *a* and *b*, be led off to a suitable instrument, *e.g.* a string galvanometer, each stimulus applied at *x* gives rise to a *diphasic* electrical variation. When the muscle is stimulated at *x*, a contraction wave commences which travels down the muscle through *a* to *b*. The electrical investigation of the muscle shows that excitation of *x* arouses an electrical change which also passes down the muscle at the same rate as the mechanical change, which it slightly precedes. On leading off from *a* and *b* there is a latent period between the stimulus and the first change, representing the time taken for the change to travel from *x* to *a*. When the change reaches *a*, this becomes the seat of a depolarisation, the membranes previously positive becoming discharged. Hence the current would pass in the outer circuit from *b*, which is still positive, to *a*. We may say therefore that *a* is negative to *b*. A fraction of a second later the excitatory change has passed on to *b* and has died away at *a*. Now *b* is negative to *a*, and the current therefore passes in the opposite direction. Between *a* and *b*, therefore, there is a diphasic current, the first phase representing negativity of *a* to *b*, and the second phase negativity of *b* to *a*. *A diphasic change is thus also a sign of a propagated change.* Every excitation of a normal muscle gives rise to a diphasic variation of such a direction that the point stimulated first becomes negative to all other points of the muscle, and this "negativity," to use a loose but convenient expression, passes as a wave down the muscle, preceding the wave of contraction and travelling at the same rate (Fig. 47).

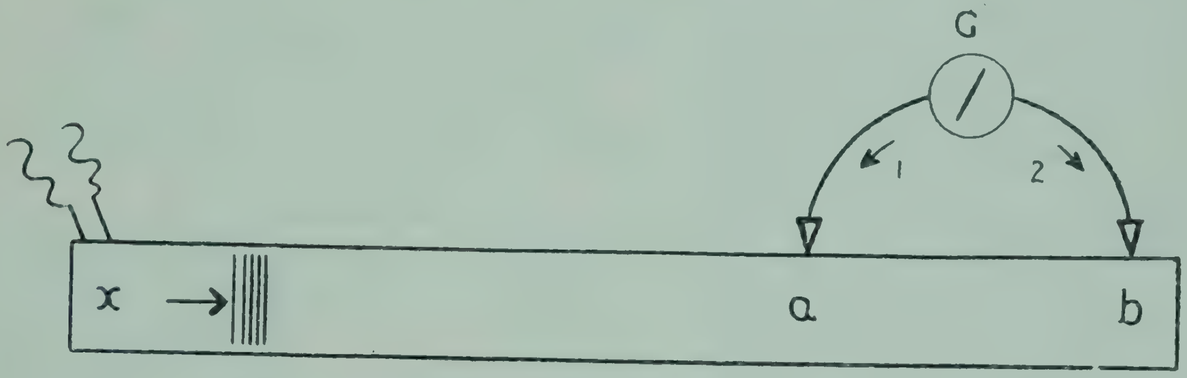


FIG. 46. Diagram showing Diphasic Variation of uninjured Muscle.

If one leading-off point be injured, *e.g.* at *b*, that point becomes permanently negative, and excitation fails at that point. A single stimulus applied at *x* will in this case give only a *monophasic* variation, or *spike*, in which *a* is relatively negative, *i.e.* less positive than previously, compared with *b*.

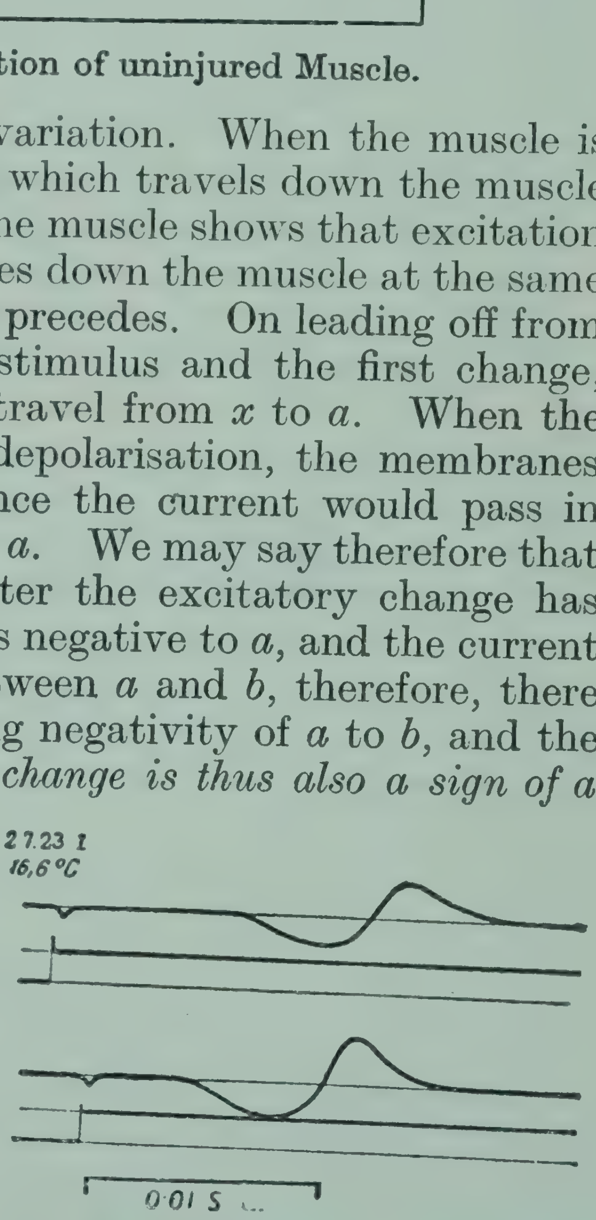


FIG. 47. Diphasic action potentials in gastrocnemius muscle of frog, recorded by an oscillograph. Below each tracing is the signal line showing stimulation of the sciatic nerve. The upper tracing shows stimulation of central end of the nerve, the lower one of the peripheral end. The time-line below indicates 0·01 sec. (ROSENBERG. *Pfluger's Arch.*, 1929, **223**, 120.)

When we study the time relation of the electrical variation ensuing on a single stimulus, we find that the electrical change under the electrodes begins at the moment that the stimulus is applied. It takes for frog muscle about 0·0025 second to attain its culminating point. At this point the mechanical change or contraction of the muscle begins. These time relations vary with the temperature of the muscle. We have already seen that the effect of lowering the temperature is to increase the latent period of the contraction. In the same way it slows the rise of the electrical change and the rate of propagation of the wave of electrical change. We are therefore justified in regarding the electrical change as an index to the ionic changes evoked in the muscle as the direct result of the stimulus. The material change in form of each contracting unit is secondary to these changes. As the result of stimulation, depolarisation is aroused at the point of excitation, and travels thence along the muscle fibre at a rate of about 3 metres per second, *i.e.* the same rate as that of the following wave of mechanical change and, like this, varying with the temperature.

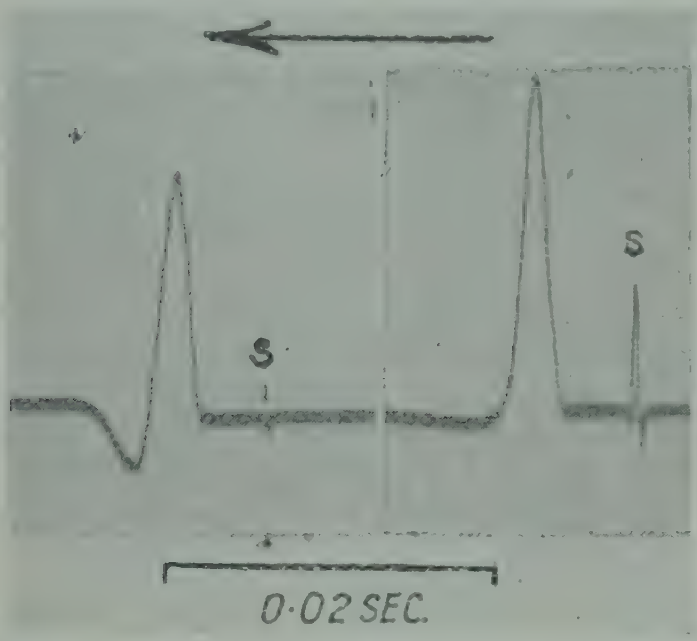


FIG. 48. Diphasic (left) and monophasic (right) action potentials in non-medullated nerve of Maia, recorded by cathode ray oscillograph. Point of stimulation shown by S. Read from right to left. Time indication = 0·02 sec. (Tracing by Dr. J. Yule Bogue.)

Under no conditions, however, may an excitatory condition be propagated without the presence of a contraction, though often this may be so small as to be difficult of detection. When a muscle is so treated as to diminish its mechanical response, *e.g.* by soaking in water, the electrical response is also reduced, and the two finally vanish together when the muscle passes into " water rigor."

The phenomena of diphasic and monophasic action potentials are shown with even greater clarity in nerve than in muscle. If two uninjured points on a nerve trunk are led off to an oscillograph, a diphasic potential wave can be recorded at the passage of each excitation applied to the nerve, and indicates the transmission along the nerve of a nerve impulse. If one of the points led off has been injured, *e.g.* by burning with a hot wire, then each nerve impulse passing the other, uninjured, point occasions a monophasic action potential (Fig. 48).

Somewhat similar tracings can be obtained from almost any tissue when excited.

For further details of nerve action potentials, see Chapter XII.

THE MECHANICAL CHANGES THAT A MUSCLE UNDERGOES WHEN IT CONTRACTS

By " contraction," in the physiological sense, we mean the state, resulting from an adequate stimulus, which would, if unopposed, lead to shortening of the muscle. If the state is opposed there may be no shortening, but even lengthening of the muscle.

If a skeletal muscle be stimulated by a single induction shock either directly or by the intermediation of its nerve, it responds by a single short, sharp contraction, followed immediately by a relaxation. The amount of shortening in contraction obviously depends on the length of the fibres, which can usually shorten by, at most, about half their resting length ; so the sartorius, with long fibres, shortens by about 50 per cent., but rectus femoris, with short fibres, only by 23 per cent. It is equally obvious that the force exerted by a muscle depends on the number of fibres present in a cross section, and on their direction with reference to the tendon or bone to be moved. In volume the muscle undergoes a very slight, and for most purposes negligible, shrinkage,* but each muscle fibre, and the whole muscle, becomes shorter and thicker.

Isotonic Contraction. If a weight be hung on to the tendon of the muscle, the muscle during contraction may raise the weight and thus perform mechanical work. Under these conditions the muscle supports the weight all the time, and on contraction lifts it. Such a contraction is called *isotonic.* To determine the time relations of the contraction it is necessary to employ the graphic method.

In order to record the shortening of the frog's gastrocnemius, the muscle is excised together with a portion of the femur to which it is attached, and often with the whole length of the sciatic nerve. The femur is clamped firmly, and the tendo Achillis attached by a thread to a light lever, free to move round an axis at one end. The point of this lever is armed with a writing point, which just touches the blackened surface of a piece of glazed paper stretched round a cylinder (drum) which can be made to rotate at a constant speed.

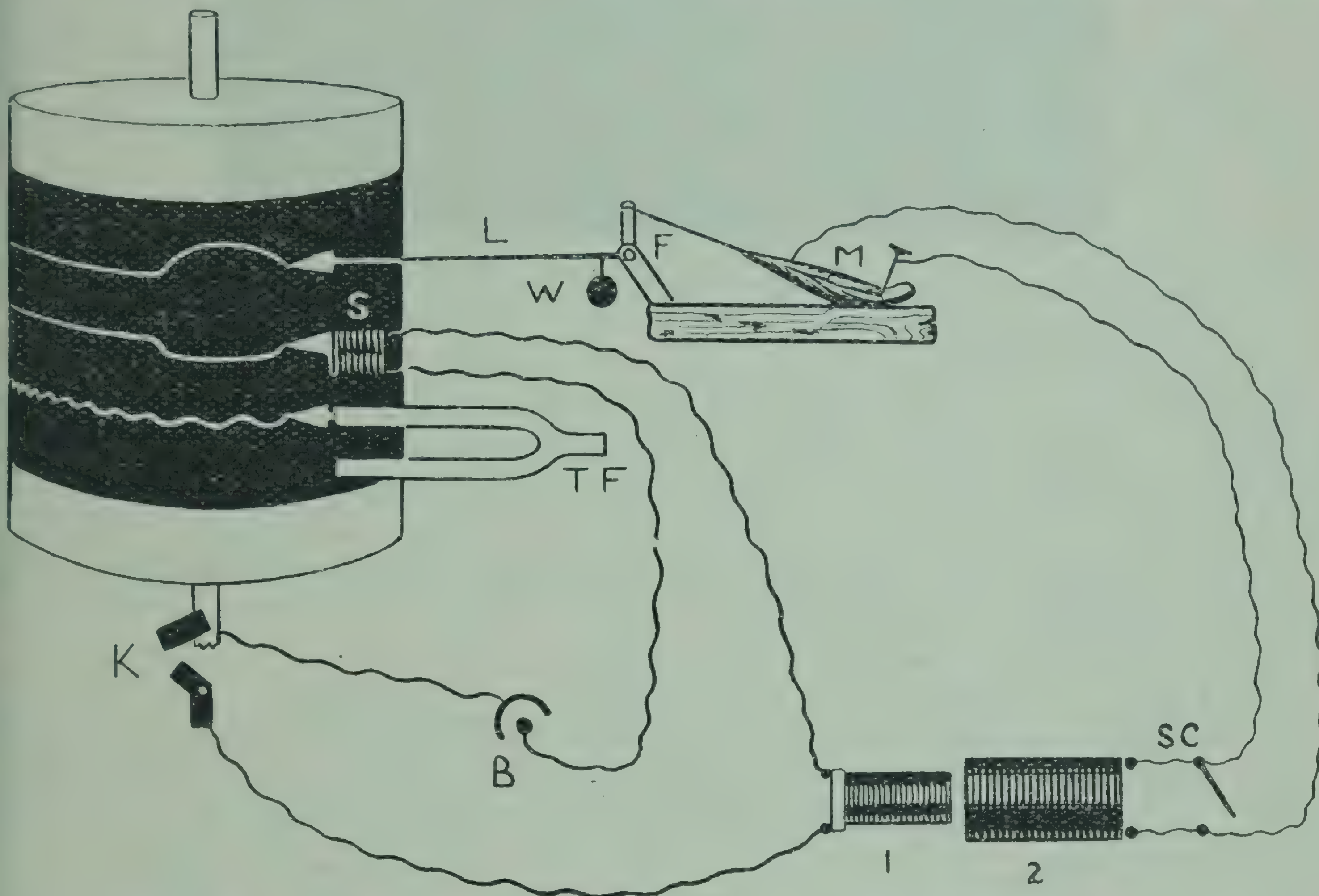


FIG. 49. Diagram of Apparatus for recording simple Muscle Twitch.

If the drum is moving, the point draws a horizontal white line or abscissa on the smoked paper.

If a single induction shock be sent through the muscle, M (Fig. 49), or nerve, the lever, L, turning on the fulcrum, F, and carrying the weight, W, is jerked up, falling again almost directly, and a curve is drawn like that shown in Fig. 50.

In all such graphic records we should have also—

(1) *A time record.* This is furnished by means of a tuning fork, T.F., vibrating at a suitable rate, *e.g.* 100 times per sec.

(2) *A record of the exact point at which the nerve or muscle is stimulated.* This may be obtained by means of an electro-magnetic signal, S, interposed in the circuit, and arranged to write on the revolving drum.

In Fig. 50 the upper line is the curve drawn by the lever as the muscle contracts; the small upright line *a* shows the point at which the muscle was stimulated; and the lower line is the tracing of the chronograph, every vibration representing $\frac{1}{100}$ of a second.

The graphic record may, in consequence of instrumental inertia, be a very inaccurate reproduction of the true events occurring in the muscle itself. When the muscle begins to contract, it imparts a very rapid movement to the lever which therefore tends to overshoot the mark and deform the curve. This source of error may be reduced by making the lever as light as possible, and hanging the extending weight in close proximity to the fulcrum of the lever, as shown in Fig. 49. Since the energy of a moving mass is proportional to the square of the velocity ($= \frac{1}{2} mv^2$) and the tension due to the weight as well as the velocity imparted to the weight on contraction is directly proportional to the distance of the weight from the fulcrum, it follows that it is better to load the muscle with 40 grammes 1 millimetre from the axis than with 1 gramme 40 millimetres from the axis, though the tensio put on the muscle will be the same in both cases.

---

* MEYERHOF and MÖHLE. *Bioch. Ztschr.*, 1933, **260**, 544, 469; **261**, 252.

In the first case the energy of the moving mass will be proportional to $\frac{40 \times (1)}{2} = 20$ and in the second to $\frac{1 \times (40)^2}{2} = 800$, and it is this energy which determines the overshootin of the lever and the deformation of the curve.

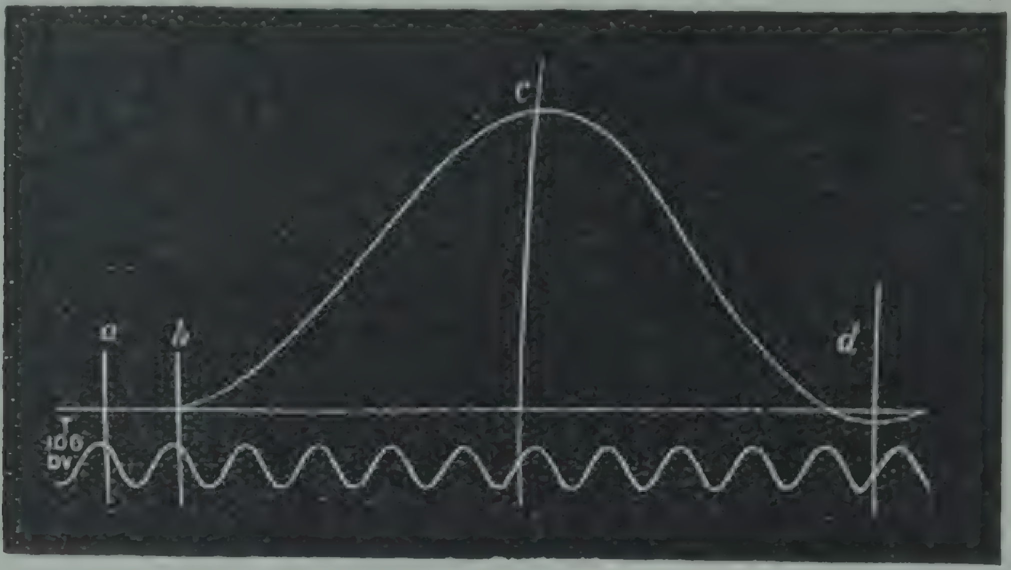


FIG. 50. Curve of simple Muscle Twitch. *a*, point of stimulation ; *ab*, latent period ; *bc*, period of contraction ; *cd*, period of relaxation. Time = $\frac{1}{100}$ sec. (After STIRLING.)

In the method just described, the lever follows the muscle in its movement, the tension on the muscle remains nearly the same throughout, and hence the term *isotonic*. An isotonic contraction or twitch, such as that in Fig. 50, produced by a single stimulus, consists of three evident phases :

(1) A phase during which no apparent change takes place in the muscle. This is called the *latent period.*

(2) A phase of shortening or *contraction.*

(3) A phase of *relaxation* or return to the original length.

The small curve seen after the main curve is due to elastic vibrations of the lever. The latent period for the frog gastrocnemius as usually recorded at 18° C. occupies about 10 msec., the phase of shortening 40 msec. and the relaxation 50 msec.

Thus a single muscle twitch in a frog is completed in about $\frac{1}{10}$ second. It must be remembered, however, that this figure is only approximate, and varies with the temperature of the muscle and its condition, being much longer in a fatigued muscle. It depends also on the animal from which the muscle is taken, and on the particular muscle. In mammals it is not necessarily shorter than in the frog. Thus in the cat, the period of actual *contraction* is, soleus 100 msec., gastrocnemius 40 msec., internal rectus of eye 8 msec. In the tortoise a single twitch may last from half a second to five seconds, according to temperature. In the wing muscle of a wasp the duration of a single contraction may be only 5 msec.

In the living body the latent period is considerably increased by reason of the inertia of the parts moved. But it is of considerable theoretical interest. When measured as above, there is also instrumental inertia, so that the true latent period must always be shorter than is apparent from a curve obtained by the foregoing method. The viscosity and extensibility of the muscle must also prolong this apparent latent period, since the first effect of contraction of any part of the muscle will be to stretch the adjacent part, and only later to move the tendon to which the lever is attached. If we have a weight suspended by a piece of elastic, it will not follow a pull exactly, but will lag behind, the first part of the pull being occupied with stretching the india-rubber, and only when this is stretched to a certain degree will the weight begin to rise.

It is possible to obviate instrumental inertia by employing photographic methods for the record. Such experiments were carried out long ago by Burdon-Sanderson. But more recent observations have brought fresh facts of interest to theory regarding the latent period.

According to Einthoven[*] and also Roos[†] the interval between the moment of stimulation and the onset of mechanical changes is less than 0·4 msec., if inertia is eliminated

[*] EINTHOVEN, Harvey Lectures, 1924–1925, p. 111.

[†] ROOS. *J. Physiol.*, 1932, **74**, 17.

as much as possible, but most observers find it to lie between 3 and 20 msec., according to emperature, in frog sartorius muscle.

If very accurate records are made of isometric contractions, the first mechanical change n frog sartorius at 0° C. is a *relaxation* * about 7·5 msec. after the stimulus : it is only about $\frac{1}{1000}$th as great as the subsequent increase of tension, which begins (at 0° C.) between 11 and 20 msec. from the stimulus (Fig. 51) ; the degree of initial relaxation is so slight as to have been missed previously.

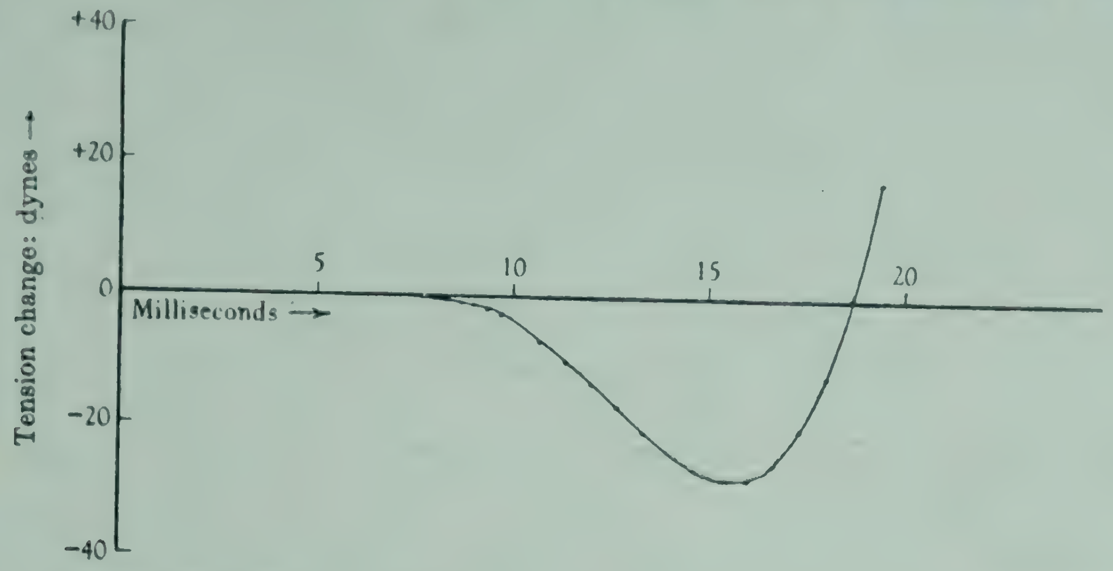


Fig. 51. Change of tension vertically against time after shock horizontally in isometric contraction of a frog sartorius at 0° C. The small initial relaxation is shown, and commencement of the tension increase. (Abbott and Ritchie, *Proc. Phys. Soc.: J. Phys.*, 1947, **107,** 5.)

The relaxation after contraction is helped by a moderate load, and in a normal condition is complete. It is probably not active but is a passive effect of extension. This may be shown by allowing a muscle to contract while floating on mercury when the subsequent lengthening on relaxation is incomplete.

Isometric Contraction. It is sometimes useful to be able to record the development of tension in the active muscle when a change in length is prevented y fixing its ends immovably. For this purpose the muscle is allowed to pull gainst a strong spring, the movements of which are magnified. Thus the hortening of the muscle is almost entirely prevented, but the increase in its

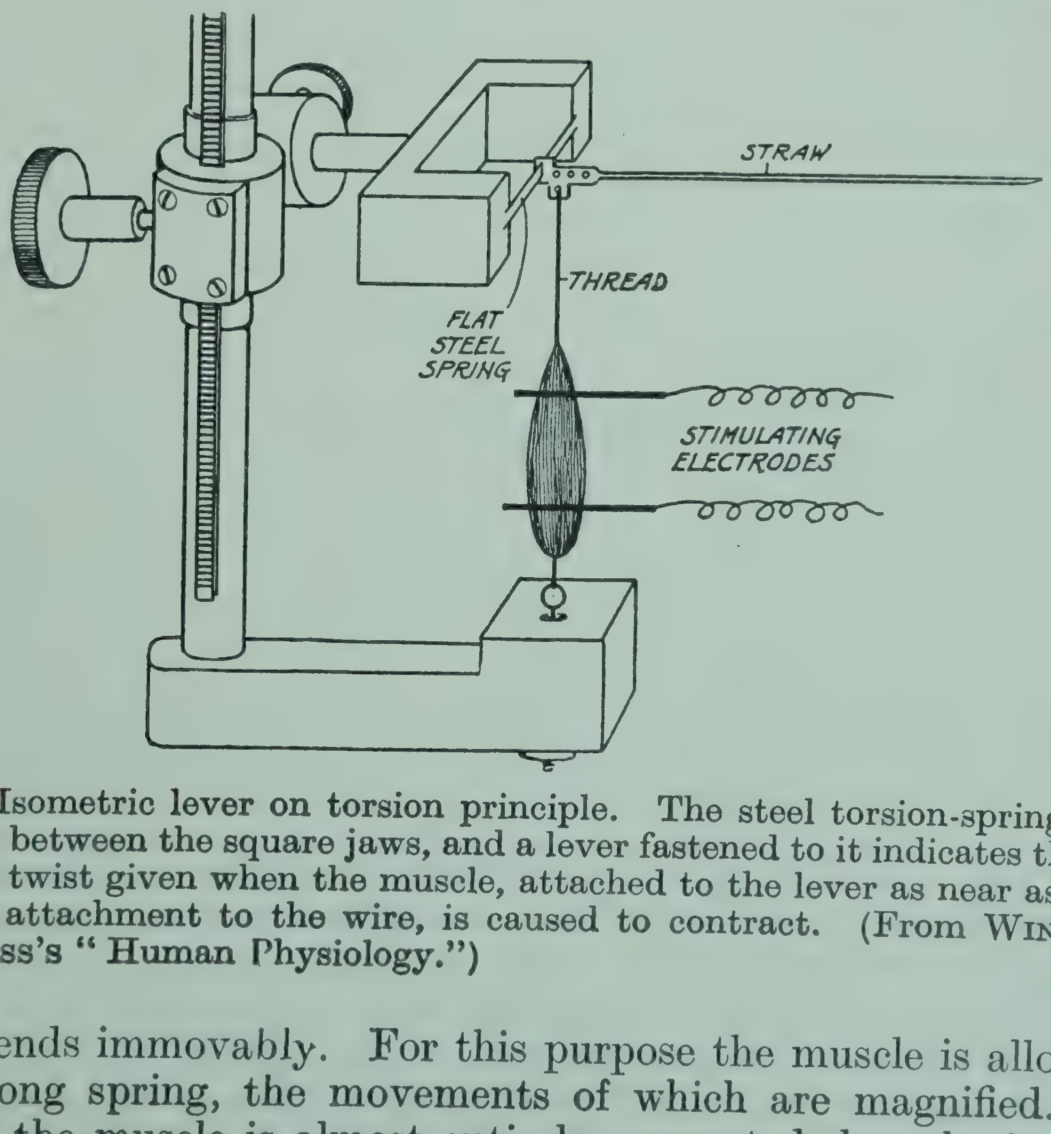


Fig. 52. Isometric lever on torsion principle. The steel torsion-spring is fixed firmly between the square jaws, and a lever fastened to it indicates the extent of the twist given when the muscle, attached to the lever as near as possible to its attachment to the wire, is caused to contract. (From Winton and Bayliss's "Human Physiology.")

---

* Sandow. *J. Cell. and Comp. Phys.*, 1941, **24,** 221.

tension causes a minute but proportionate movement of the spring, which is suitably recorded. Since in this case the length of the muscle remains approximately constant throughout, it is known as the *isometric* method. The great magnification necessary in this method necessitates the use of instruments having low inertias and very high natural vibration frequencies (about 1,000–2,000 per second), as otherwise the record is deformed by the vibrations of the spring. The simplest arrangements are on the torsion-wire principle, one type of which is shown in Fig. 52. A piezoelectric device is probably more precise.

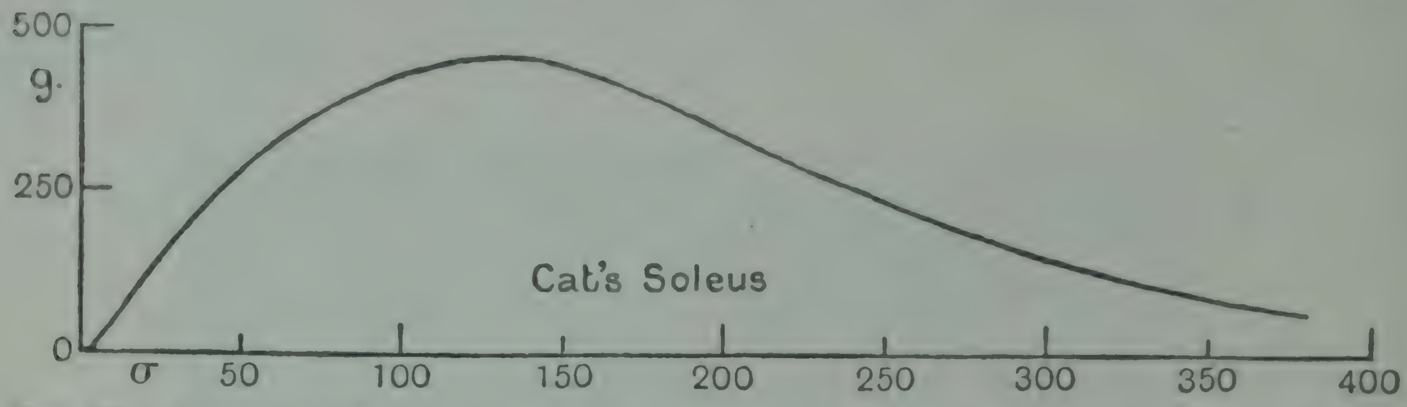


Fig. 53. Isometric Contraction of Cat's Soleus Muscle. Abscissæ = time in msec. Ordinates = grammes tension. (Cooper and Eccles, *Journ. Physiol.*)

The *isometric twitch* shows that tension change with time has a form roughly similar to that of the isotonic contraction relating length change with time, but differing from it in detail. The rise of tension commences very abruptly about 1 msec. after the beginning of the electrical response (mammal), the curve showing at first an upward concavity, soon followed by an upward convexity. After reaching the summit, there is an upwardly concave passive relaxation (Fig. 53). The duration of the contraction process is given by the interval between the onset of electrical response and the summit of tension production.

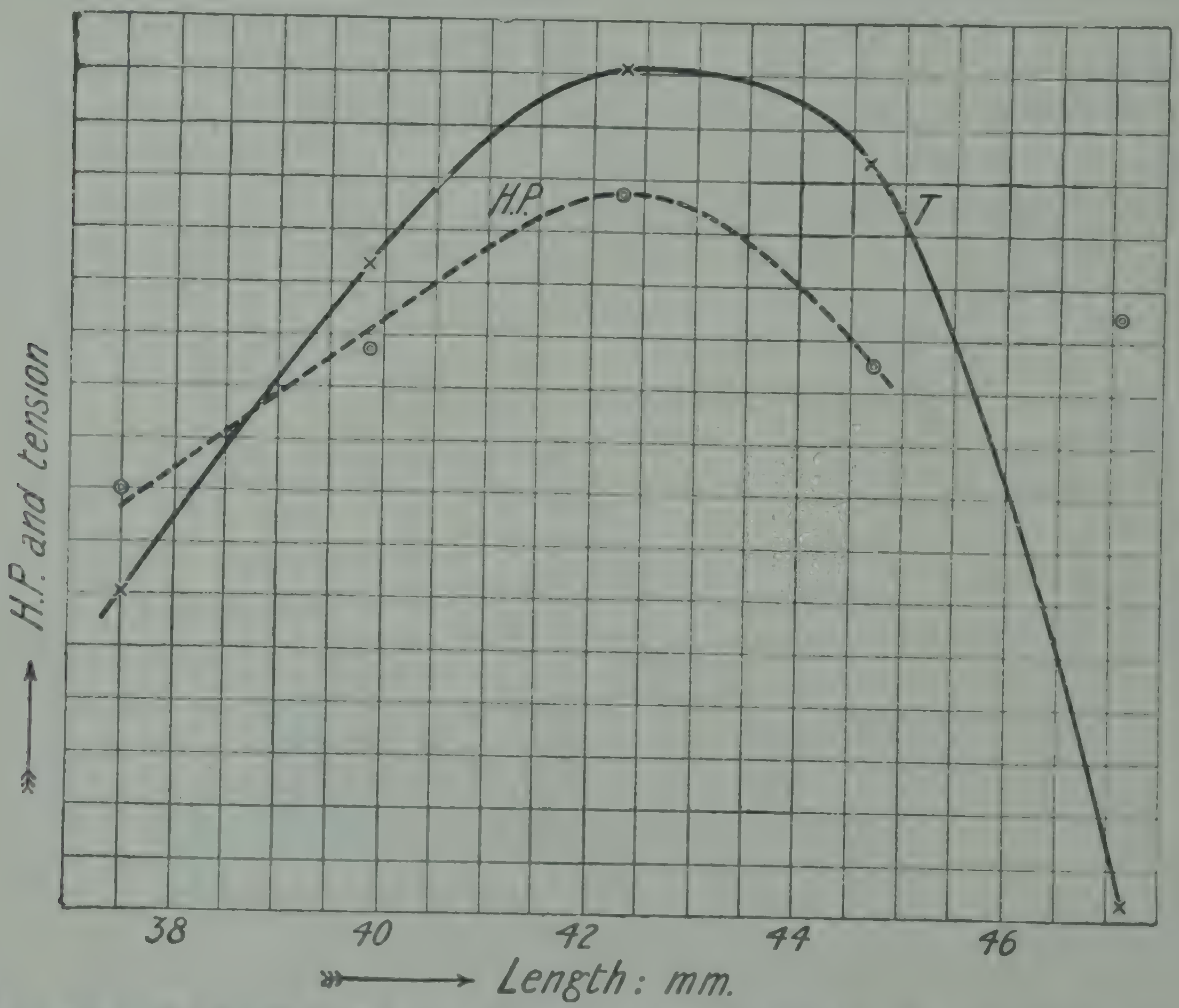


Fig. 54. Relations of heat and tension production on isometric contraction to initial length of muscle. (After Evans and Hill. *Journ. Physiol.*, **49**, 10.)

Very considerable tensions are produced in isometric contractions; several hundred grammes with a small frog's gastrocnemius. The maximum is reached sooner, but the duration of the contraction process is definitely *longer* in an sometric twitch than in an isotonic contraction ; the duration is also increased y increasing the "initial" tension on the muscle, *i.e.* by stretching the muscle efore its stimulation. Isometric contractions have been recorded in single ibres.*

The isometric contraction is a practical convenience, but is theoretically a pure abstraction, because, in order to record the tension changes, *some* shortening, though slight, must occur : also, even if approximation of the extremities of the muscle is quite prevented, internal shortenings occur during the contractile stress, and these changes introduce complications into the theoretical study of the energetics of contraction.

Effect of Initial Length of Fibre. The tension developed in a muscle contracting isometrically in response either to a single stimulus or to a short tetanus is not constant but is found to vary with the initial length of the muscle (Fig. 54). As the length of the resting muscle is increased by extending it, up to a length somewhat greater than that which it possesses in its normal extension in the body, the tension developed on contraction also increases.† Single fibres also show this property.‡

In the body the bony levers are so arranged that the muscles at their greatest length work at a maximum mechanical disadvantage which lessens continuously as the muscles shorten and approximate their points of attachment. The load on a muscle is thus diminished continuously as the muscle contracts.

MECHANICAL PROPERTIES OF MUSCLE

The operation of the musculature in bringing about skeletal movements is a question of the direct application of elementary anatomy and mechanics : with such knowledge the bodily movements are easily explained. But the causation of muscular contraction is a primary physiological problem.

So far as we know, the only function subserved by muscle is the mechanical one of contracting and so causing movement. It is therefore important to have as clear an understanding as possible of the differences in the physical properties of muscle in the resting and contracted states, and of the energy changes involved in contraction and relaxation. Yet despite much work on the subject these properties are so far by no means adequately studied, or even properly defined. This is largely because the properties of a muscle, considered as a whole, are the resultant of the properties of its different structural components, *e.g.* contractile substance proper, sarcolemma, connective tissue sheaths, etc. ; the separate properties of which must differ considerably.

Tension-length Relationship. When a resting muscle is stretched, the relation between tension and length is not in agreement with Hooke's law, which says that strain (= elongation) is proportional to the stress (= tension) producing it. Instead, we find that the successive lengthenings produced by successive equal increments of tension become less and less, so that, over a large range the tension is an exponential, instead of a linear, function of the stretch. The explanation of the departure from Hooke's law is probably to be found in the compound make-up of a muscle ; the isolated fibre, however, shows similar properties, but it has been claimed that the sarcolemma alone also does, so that the properties of the actual contractile material in this respect are unknown.

The tension-length diagram of a tetanically contracted muscle is altogether different (Fig. 55). At lengths between 60 per cent. and 100 per cent. of the normal extended length of the muscle the relation is roughly linear, as Hooke's law would lead us to expect, but at lengths over 100 per cent. the tension *falls off* with increasing length, though up to 160 per cent. this is reversible : above 160 per cent. the stretching is not fully reversible. At lengths *below* about 60 per cent. of the normal extended length the shortening is also partly irreversible, the fibres then being said to be in the delta state. These relationships mean that, as already explained, at the lengths at which muscles normally function in the body, the greater the initial length of the muscle, the greater th tension produced when the muscle contracts isometrically.

---

* Brown and Sichell. *J. Cell. and Comp. Phys.*, 1936, **8**, 315.

† Evans and Hill. *Journ. Physiol.*, 1914, **49**, 10 ; Doi. *Journ. Physiol.*, 1921, **55**, 38.

‡ Ramsey and Street. *J. Cell.and Comp. Phys.*, 1940, **15**, 11.

Energy Disposal in Shortening. When a muscle shortens the tension set up is utilised either to lift a load or accelerate a mass, to act on its own elasticity, or to overcome its own viscous resistance. Much work on these properties has led to conflicting results and it is now clear that the problem cannot be reduced to these simple terms. The difficulties are due to the complex structure of muscle and to the fact that the muscle fibres do not instantly develop the full amount of energy, and then utilise this tension to do work, etc., but *develop the energy of contraction as shortening proceeds*, and at a rate which depends on the speed and extent of shortening. If the muscle is allowed to shorten, more energy is set free than if the contraction is isometric (the Fenn effect) : if it lengthens on " contraction " (as when there is too heavy a load) then the energy supply is reduced. Consequently, any attempts to compare the viscous and elastic properties of resting with those of contracting muscle are defeated, because the properties of the contracting tissue are altered by the methods used to test those properties. In fact, the muscle is living, and therefore has the property of adjusting itself to altered conditions. This shows itself as a " stiffness " of the contracting muscle, which has sometimes been regarded as due to an alteration in its visco-elastic properties.

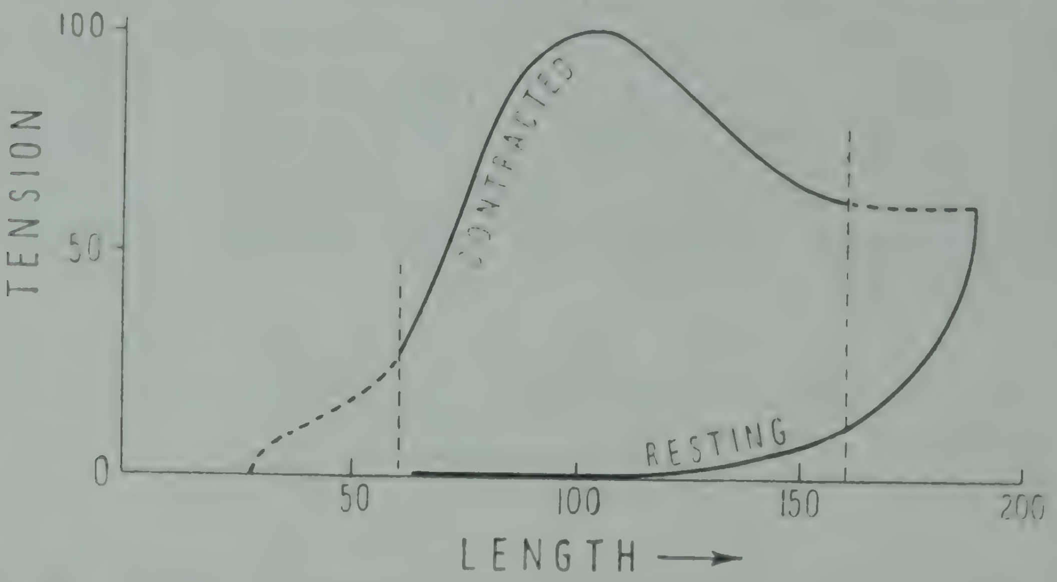


Fig. 55. Relation between tension and length in resting and fully contracted muscle.

It is for these reasons that muscle shows a decrease of force with increase in speed of shortening, and a decrease in speed of shortening with increase of load.

Effect of Temperature. When a muscle is moderately warmed, the chemical processes involved in the phase of shortening as well as those responsible for relaxation are hurried up, while the "viscosity" is diminished. The latent period therefore becomes shorter and the muscle curve steeper and shorter.

It is very often observed that in an isotonic twitch the height of the curve of contraction of the warmed muscle is greater than that obtained at ordinary temperatures. This apparent increase in height is really instrumental in origin, the quicker moving muscle jerking the lever beyond the real extent of the contraction.

When isometric contractions are recorded, it is found, at any rate in the frog, (1) that the tension developed in a single twitch diminishes slightly as the temperature is raised from 0° C. to 25° C. ; (2) that the tension developed in a tetanus lasting a sufficient time to attain its maximum is increased with rise of temperature, but not to a very marked extent ; (3) that the durations of the latent period and contractile process are greatly reduced, and the duration of relaxation somewhat reduced by rise of temperature.

If a muscle be heated gradually (without stimulation) up to about 45° C., it begins to contract slowly at about 34° C., and this contraction reaches its maximum at 45° C., at which point the muscle has entered into pronounced *rigor mortis.* This form is often called heat rigor.

Cold has the reverse effect. All the processes of muscular activity are slowed, so that the latent period and the contraction period are prolonged. It is therefore often convenient for analysis of the progress of events in contraction to use frog muscle at 0° C. Slight cooling also lowers the *threshold of excitability* of muscle, so that any form of stimulus is more effective at 5° C. than at 25° C. If a muscle be cooled for a short time to a little below zero, it loses its irritability, which returns if the muscle be gradually warmed again. Prolonged exposure to severe cold irrevocably destroys its irritability, and warming the muscle will now simply being about *rigor mortis*.

PROPAGATION OF CONTRACTION. The whole muscle does not, as a rule, contract simultaneously. When excited from its nerve the contraction begins at the end plates and spreads in both directions through the muscle. The rate of propagation of the contraction wave can be measured only by employing a curarised muscle, so as to avoid the wide spreading of the excitatory change by means of the intra-muscular nerve twigs. For this purpose a curarised sartorius muscle is stimulated at one end, and the electrical variation in the muscle recorded by two pairs of electrodes, one (*a*) placed near the exciting electrodes and the second (*b*) at the other end of the muscle. The difference between the times of commencement of the electrical variations at the two places gives the time taken by the contraction wave in travelling from *a* to *b*.

By measurements carried out in this way it is found that the rate of propagation of the contraction in frog's muscle is 3 to 4 metres per second ; in the muscle of warm-blooded animals it may amount to 6 to 12 metres. The *length of the wave* is obtained by multiplying the rate of transmission by the duration of the wave at any one point. It varies in frog's muscle between 150 and 360 millimetres. Thus the muscle fibres are much too short to accommodate the whole length of the wave, and the contraction of the whole muscle must be made up of the summated effects of the contraction wave as it passes from point to point.

The Effect of Strength of Stimulus. The "All-or-None" Rule. If a series of single break induction shocks be applied to a muscle at intervals of not less than five seconds, it will be found that when these are below a certain strength no effect at all is produced. The shocks are said to be *subliminal*. On gradually increasing the strength a point will be reached at which a small contraction will be observed (*liminal* or *threshold* stimulus). Further steady increase in the strength of stimulus will now result in a gradual increase in the power of the contractions (*submaximal stimuli*) until finally a maximum response is reached (*maximal stimulus*) ; further increase in the strength of stimuli (*supra-maximal stimuli*) cause no further increase in contraction, although the current actually used may be increased a hundred times. It might seem that this limited gradation of the muscular response according to strength of stimulus was due to a similar gradation in the response of each of the muscle fibres of which the muscle is composed. The fact is, however, that when a minimal or submaximal response is obtained, not all the fibres making up the muscle are contracting. A minimal contraction is, in fact, a contraction in which *a few* fibres of the whole muscle are stimulated. A maximal contraction is one in which *all* the fibres are stimulated. So far as concerns each individual muscle fibre every contraction is probably a maximal contraction, *i.e.* the rule of "all-or-none" which was first enunciated for heart muscle is probably true for skeletal muscle. The difference between skeletal and heart muscle lies in the fact that in the former the excitatory process does not spread from one fibre to its neighbours. If, for instance, we take a curarised sartorius and split its lower end as in Fig. 56, the stimulus applied to A causes a contraction only of the left side of the muscle, while a stimulus applied to B in the same way limited to the right side. If a piece of ventricular or auricular muscle of the frog or tortoise were treated in the same way, a stimulus applied at A would cause a contraction which would extend to B.

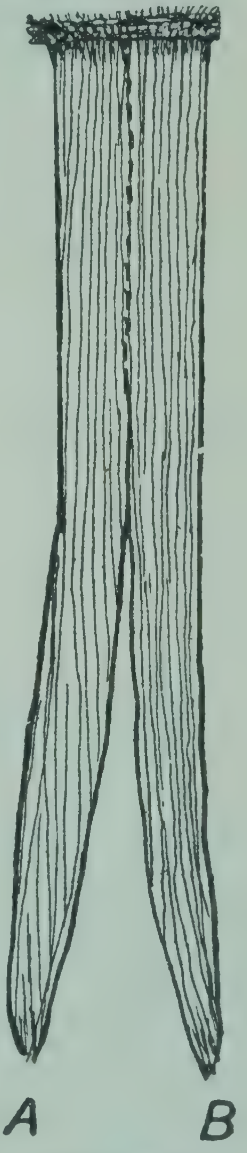


FIG. 56.

The "all-or-none" law does not mean that it is impossible to vary *by any means* the extent of the mechanical response of a fibre to stimulation, but only that such variations cannot be brought about by altering the strength of the stimulus.

The "all-or-none" character of muscular contraction was indicated by Gotch and Keith Lucas,* but the clearest demonstration of it is that given by the experiments of Pratt and Eisenberger.† These investigators were able, by the use of micro-electrodes, to apply stimuli to individual fibres, while observing these directly under the microscope, an also to take photomicrographic records of the extent of the contractions which resulted One of these records is shown in Fig. 57. It will be noticed that, as the strength of th stimulus is steadily increased, the response increases in a series of steps, each of whic was seen to correspond to the occurrence of contraction in a fresh fibre or fibres which had been unaffected by the weaker stimuli: each fibre which enters into contraction does so with maximal strength. A grading of muscular power is thus achieved by variation in the number of fibres or parts of fibres brought into play. Further investigation of the phenomenon by other workers suggests that this simple view may need to be qualified, since when very fine electrodes and weak stimuli are used, the response in their immediate neighbourhood, up to a point, is graded.* This is apparently due to the fact that weak and very localised stimulation of a single fibre causes a localised contraction which does not propagate and occupy the whole fibre.† But beyond a certain point of intensity, propagation succeeds, and then the whole fibre springs into activity, and this is all-or-none.

---

* LUCAS. *J. Physiol.*, 1909, **38**, 113.

† PRATT and EISENBERGER. *Amer. Journ. Physiol.*, 1919, **49**, 1.

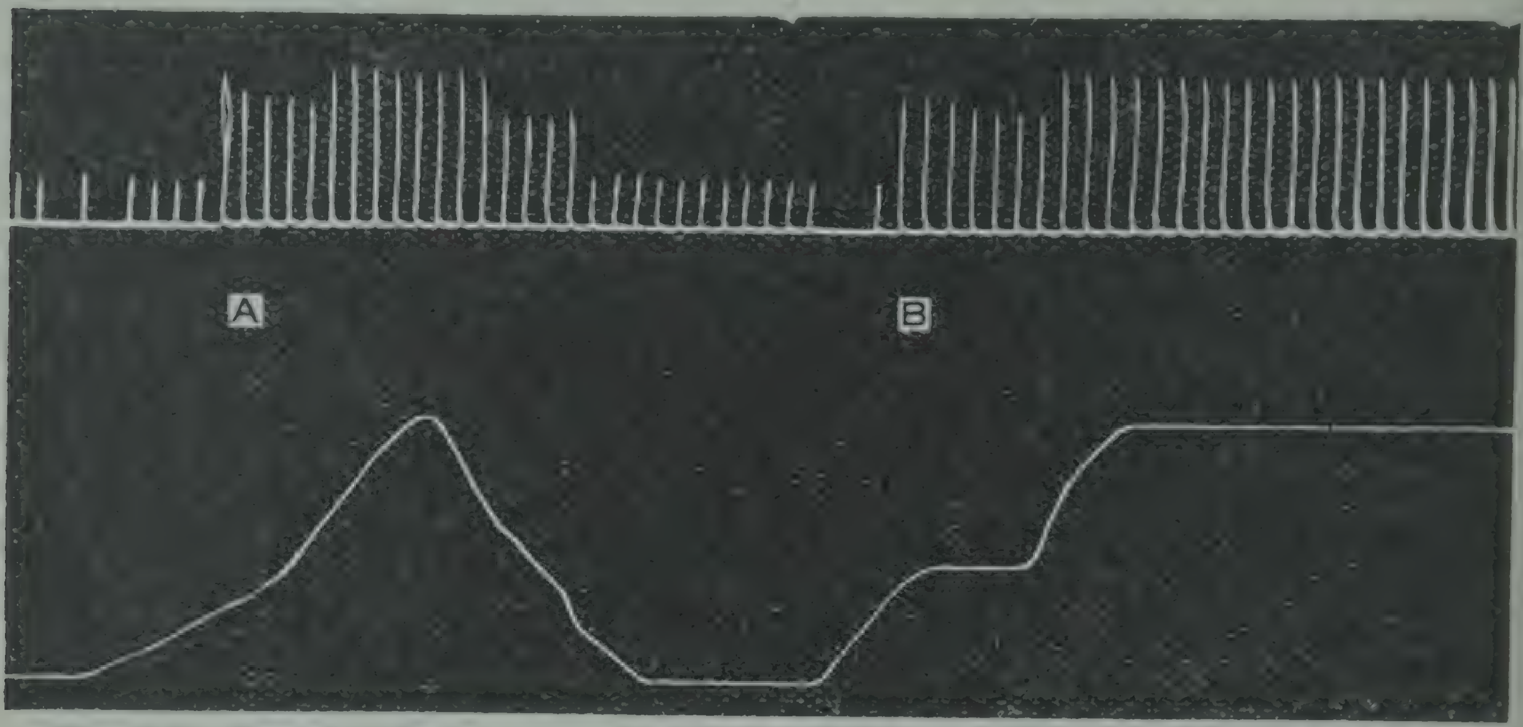


Fig. 57. Record of Contraction of Sartorius of Frog stimulated by break shocks from a " pore " electrode. The record was made by photographing the excursion of a droplet of mercury on the surface of the muscle. Upper tracing = extent of successive contractions; lower tracing = variation in intensity of stimulus. (After Pratt and Eisenberger.)

For Reference

Ritchie (1932). The All-or-None Principle. *Biolog. Rev.*, **7**, 336.

THE REPETITION OF STIMULI

Summation of Contractions and Tetanus. The response of a single muscle fibre to a single shock, whether measured by the isotonic or the isometric method, *i.e.* as shortening or as tension, is independent of the strength of stimulus and varies only with the length of the fibre. If, however, a second shock is sent in during the rise of the excitatory condition, a further evolution of energy is possible, and the effect is still further increased by putting a series of stimuli into the muscle, or its attached nerve, before the development of the contractile stress due to each previous stimulus has reached its maximum. If two shocks at intervals of, say, 20 msec. be sent into a cat gastrocnemius muscle, the response will be greater than that produced by one shock (Fig. 58). The effects of the two stimuli are summated; the resulting contraction may be less than twice the size of the simple twitch, or it may be more.‡ Thus, in the ocular muscle of the cat, stimulated via its nerve, summation begins at about

---

* v. Gelfan and Gerard. *Amer. J. Physiol.*, 1930, **95**, 412.
† Pratt. *Amer. J. Physiol.*, 1930, **93**, 680.
‡ Cooper and Eccles. *Journ. Physiol.*, 1930, **69**, 377.

0·5 msec. interval (the R.P.), and is maximal at 2·5 msec., and is then about 2–3 times the response due to a single stimulus.* With 5 (or more) msec. separation the two responses separate out. (This phenomenon of *summation of contractions* must be clearly distinguished from that of summation of stimuli referred to previously.)

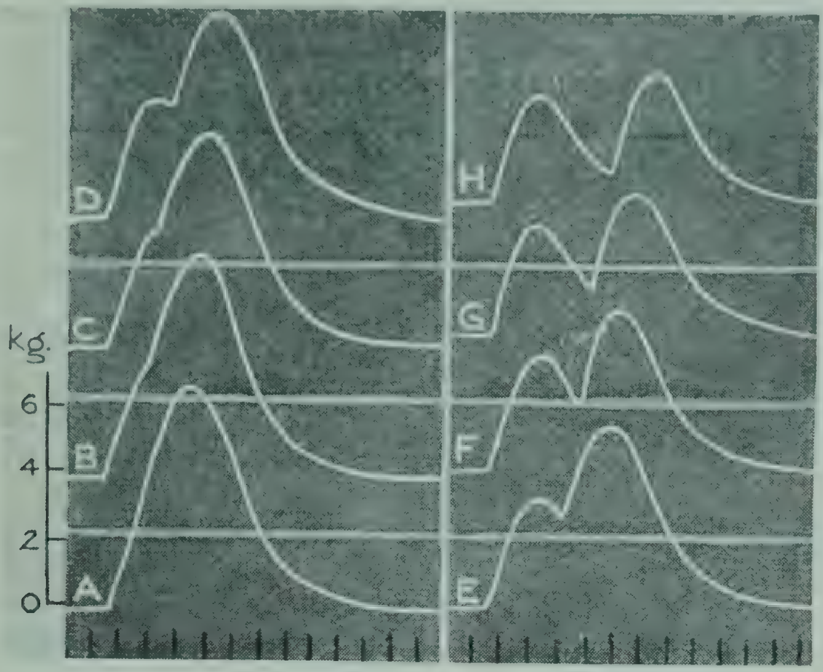


FIG. 58. Isometric Responses of Median Head of Gastrocnemius of Cat, *in situ*, to two maximal stimuli to peroneal nerve.

The intervals between the two shocks were lengthened in order from A (shortest interval) to H (longest interval). The times between the first and second shocks were, in msec.: A, 24; B, 32; C, 40; D, 48; E, 57; F, 69; G, 77; H, 88. Time intervals marked are 20 msec. (After COOPER and ECCLES, *Journ. Physiol.*, 1930, **69**, **377**.)

If a series of shocks be sent in, the excitatory condition is maintained, so that instead of a simple muscle twitch, rising to a maximum and then falling,

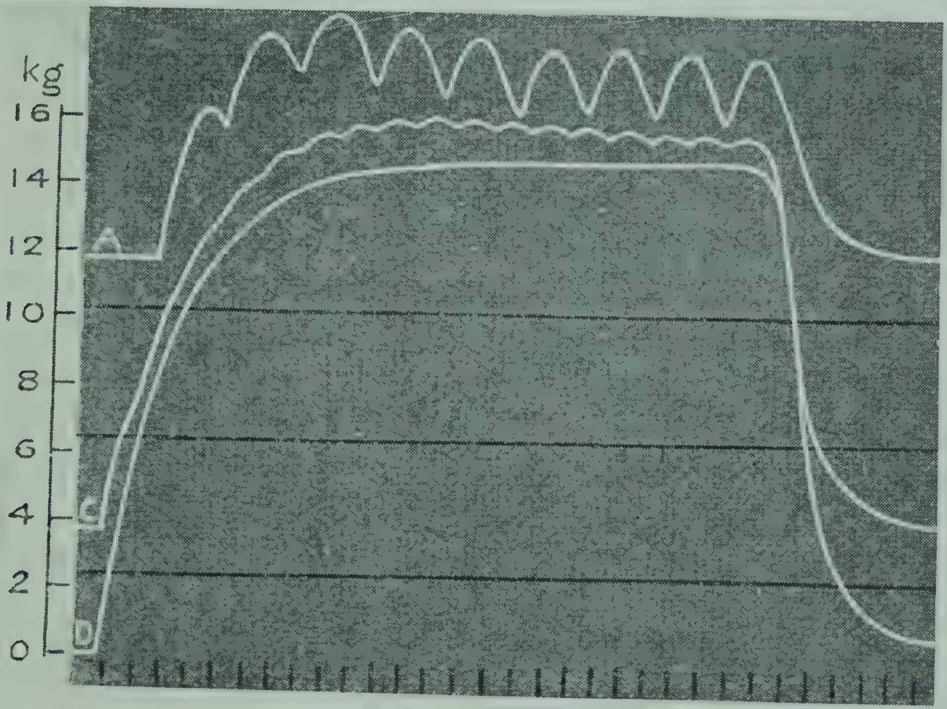


FIG. 59. Isometric Responses of Median Head of Cat's Gastrocnemius to rapidly interrupted Repetitive Stimulation of Popliteal Nerve. A at 19; C at 35; D at 115, shocks per second. (COOPER and ECCLES.)

he muscle lever rises to a given point (Figs. 59 and 60), which in the muscle ontracting isometrically may be four times that due to a single stimulus and hen remains at this height during the continuance of the repeated excitations. he state of apparently continuous contraction evoked in a muscle by repeated hocks is called a *tetanus* (Fig. 59, D). If the stimuli are not rapid enough to ive complete fusion, a partially maintained contraction, called incomplete tetanus, or *clonus*, results (Fig. 59, A, C). The rate necessary to give tetanu is higher the shorter the contraction time of the muscle, and therefore varie from muscle to muscle. If the muscle be allowed to contract isotonically the sustained shortening produced by a series of stimuli may, with a heav load, be three or four times as great as that produced by a single stimulus.

---

* BROWN and HARVEY. *J. Physiol.*, 1941, **99**, 379.

When, as under the conditions of ordinary life, the muscle contracts in response t impulses transmitted to it by its motor nerve, the conditions are not quite the same as whe the muscle is stimulated directly. Each motor nerve fibre supplies many muscle fibres often about 150.* The unit of response i therefore a nerve fibre with all the muscl fibres which it supplies; this is called a *moto unit.*

When a motor nerve is stimulated elec trically with a maximal shock, we may suppose that all its motor units are thrown into action. If a second shock follows quickly on the first, and both are maximal, we get summation, as explained above. The same phenomenon is seen when the stimuli are submaximal.

This type of summation may be called *wave-summation,* because it is due to the fusion of successive waves of contraction in the same fibres. This summation is not algebraical, *i.e.* the second contraction does not exactly double the shortening brought about by the first, nor does it double the duration; the effect of the second stimulus is greater the earlier it falls during the first response.

When the nerve is stimulated with a strong shock, it causes a larger contraction than if excited by a weak stimulus. This is an example of another type of summation called "*multi-fibre*" or *quantal* summation, which is due to an increase in the number of motor units responding, each fibre giving an all-or-none contraction. Thus in submaximal responses, as the strength of stimulation is increased, more and more fibres respond (synchronously). Or, again, when a mixed muscle containing fibres of different contraction durations (*e.g.* red and white fibres) is stimulated a summated contraction combining the features of both will result. Quantal summation is algebraical in character.

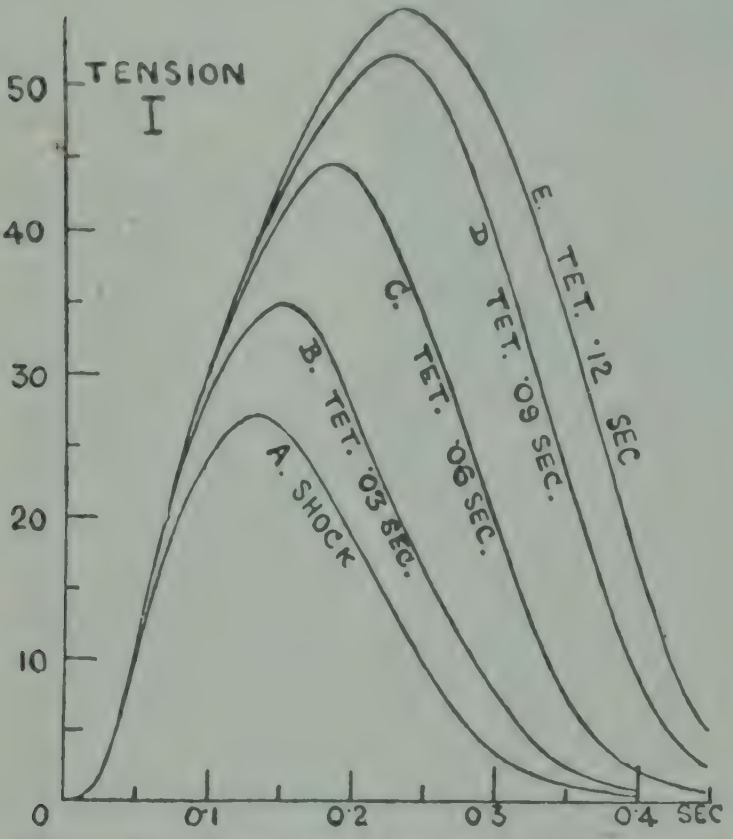


FIG. 60 Isometric Contraction Tetani showing continued rise of tension with prolongation of the excitation.

REFRACTORY PERIOD. If the interval between two stimuli sent into a curarised muscle be successively shortened in a series of observations, we finally arrive at a point at which summation of contraction is no longer apparent, *i.e.* the effect of the two stimuli is no greater than the effect of a single stimulus. This means that for a *very* short period of time after stimulation a muscle is inexcitable to a second stimulus. The period during which it is inexcitable is known as the refractory period and amounts in frog's skeletal muscle to about 5 msec. It is shorter in mammalian muscle (less than 2 msec.), and in all types is greatly increased by reduction of temperature.† The same phenomenon is a common property of excitable tissues generally.

RELATION BETWEEN MECHANICAL AND ELECTRICAL RESPONSE

We have seen on p. 119 that excitation of muscle leads to the developmen of an action potential, the appearance of which may be accepted as an indicatio that an adequate stimulus has been applied. When the time relations of th stimulus, the action potential and the mechanical response are compared it i seen that the action potential commences almost precisely when the stimulu is applied, but that it is practically over before the mechanical response begins

---

* CLARK. *Journ. Physiol.,* 1930, **70,** 18P; CREED *et al.* Reflex Activity of the Spina Cord. (Oxford.) 1932.

† BAZETT. *Journ. Physiol.,* 1907, **36,** 414; KEITH LUCAS. *Journ. Physiol.,* 1910, **39** 331.

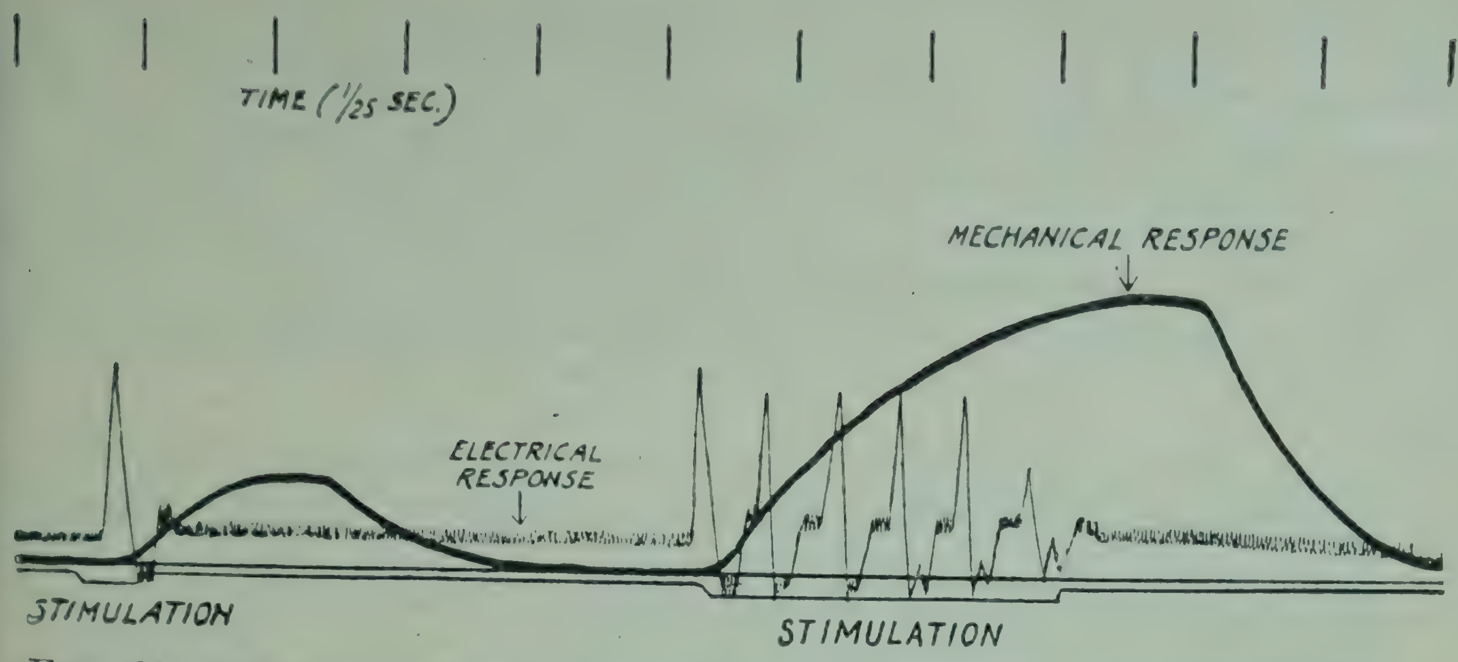


FIG. 61. Relation between Mechanical and Electrical Responses of frog's gastrocnemius *in vivo*, to (1) a single shock, (2) a rapid succession of six shocks. (After FULTON.)

It is further seen that when stimuli are repeated at short intervals, so as to result in a smooth tetanus summation, the action potentials do not so fuse, but each one separately appears after each stimulus (Fig. 61).

THE TONUS OF MUSCLE

The muscles in the normal body differ from excised muscles in being frequently in a condition of slight tension, which is called tone or tonus. This is due to reflex activity, since it is lost if the spinal cord is destroyed or if the motor nerves to the muscles are severed: the denervated muscle is toneless and inert.

If the tendon of a muscle *in situ* be cut and pulled, so as to stretch the muscle, the muscle pulls back against the extending force, *i.e.* its state of tonus, or tonic contraction is increased. When there is no pull upon its tendon there is little tonus, but as soon as the tendon is pulled upon the tone augments and increases with increase of the tension. This is called the "*stretch-reflex*" and is more fully dealt with under the central nervous system. If the motor nerve is cut the tonus is abolished, so it is obviously a state transmitted to the muscle through its motor nerve. But it can be shown that the slight constant contraction of tonus is not due to a slight continued shortening of all the muscle-fibres, but rather to their working, a few motor units at a time, in relays. These contractions, and the action potentials that precede them, can be recorded in single motor units if sufficiently precise methods are used (Fig. 62). It is then seen that at the commencement of the stretch-reflex the tone is stepped up by wave-summation, resulting from double discharges, but is then maintained by rhythmic single discharges, producing slight contractions. For this reason, tonus is not followed by fatigue and may be, and is, kept up steadily for hours on end if the conditions are appropriate.

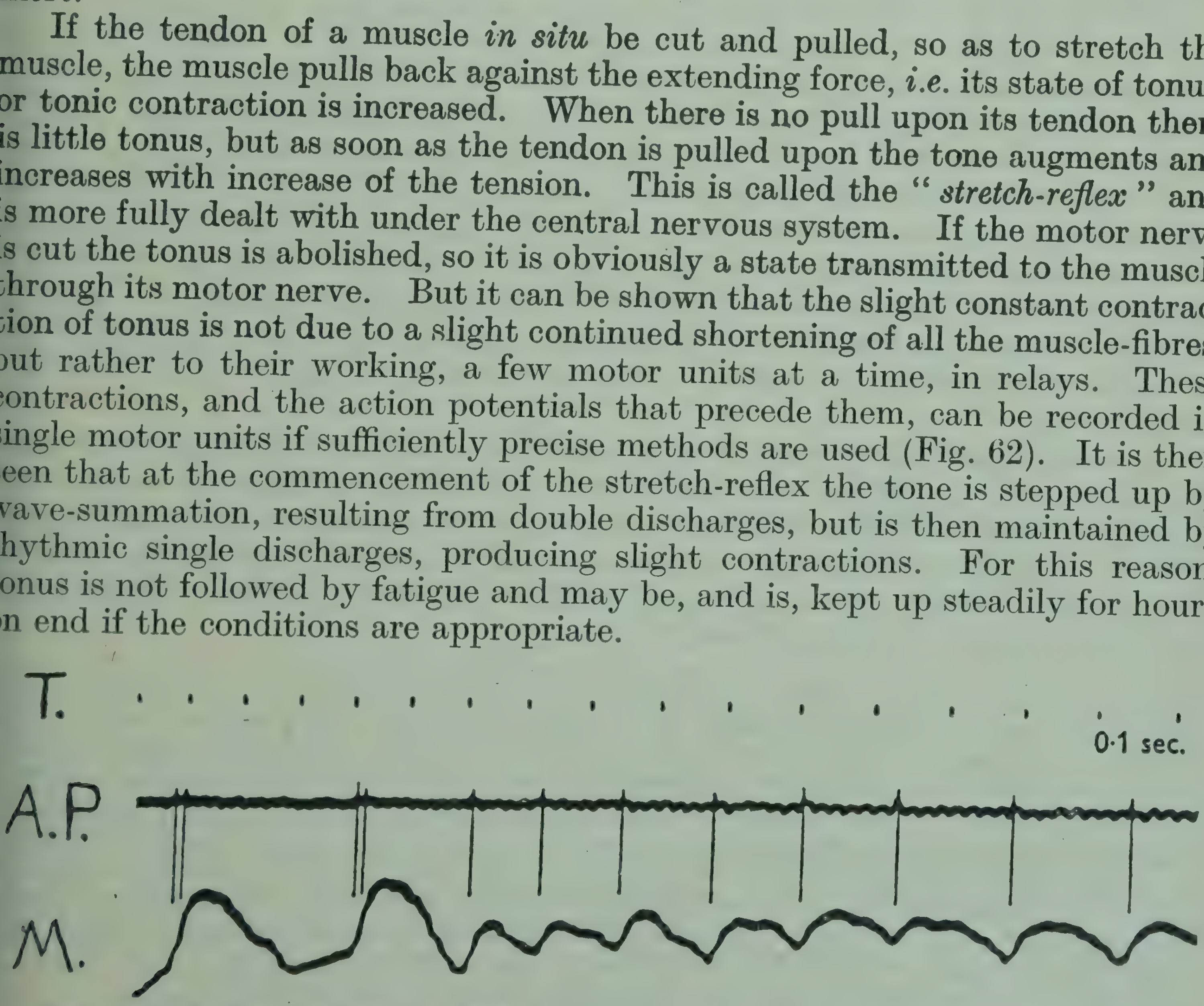


FIG. 62. Action potentials (A.P.) and movements (M.) in single motor unit of *crureus* muscle of decerebrate cat during a stretch-reflex. (GORDON and HOLBOURN. *Proc. Physiol. Soc.*, 1948, **107**, 18.P.)

Muscle tonus is also lost if the afferent nerves from the muscle and its tendon are divided, even though the motor nerves remain intact. It follows from this that the reflex involved in the tonus arises in the muscle itself. The muscle is supplied with afferent end-organs in the muscle spindles, and the nerve endings in the tendons, and when the muscle is pulled upon these originate a stream of nervous messages which passes to the spinal cord and there initiates the discharge along the motor nerves to which the tonus is due ; the stronger the pull upon the muscle the more frequent the nerve impulses which are sent from it.

Tonus is of great importance in all our muscular responses ; our posture is largely maintained by its action, and all our voluntary movements take place on a background of general tonus. In diseased conditions, where it is lost, the clumsy movements bear witness to the disability occasioned by its absence.

For Reference

Bremer (1932). Le Tonus Musculaire. *Ergebn. d. Physiol.*, **34**, 678.
Cobb (1925). The Tonus of Skeletal Muscle. *Phys. Rev.*, **5**, 518.

VOLUNTARY CONTRACTION

Under normal circumstances the contraction of skeletal muscles is brought about either reflexly, or in response to some stimulus descending from the cerebral cortex, the so-called "voluntary contraction." These contractions may have a duration of almost any extent from about 0·1 sec., but nearly all voluntary contractions last considerably longer than that.

Since we know no certain means of extending contractions to any given length except by means of summation, it is natural that physiologists have regarded voluntary contractions as similar to the artificial tetanus produced by a summation of effect, either as wave summation or as quantal summation. In the belief that the response of each muscle fibre is invariably all or none in character, it must be supposed that the delicate gradation shown in the power of voluntary contractions is due to the participation of a variable number of motor units, the maximal force being exerted if and when all the fibres contract. Smoothness of contraction could be ensured by either mode of summation. Thus by a wave summation the same set of fibres would remain in a state of tetanic contraction throughout, whereas in quantal summation different sets of fibres would work "in shifts," each fibre giving a short tetanus or even a single twitch.

It was formerly supposed, in the belief that the response was synchronous in all fibres, that definite information as to the nature of voluntary contraction would be obtainable by a study of the electrical variations or of the mechanical records ; but the accumulated evidence of quantal summation of asynchronous contractions renders such investigations of doubtful value unless single fibres, or at most only a few fibres, could be studied.

The impulse, which normally travels down the motor nerve from the anterior cornual cell to the muscle, is discontinuous. Gasser and Newcomer [*] investigated the electrical changes accompanying contraction of the diaphragm, and also those in the phrenic nerve. They found that each contraction of the muscle is accompanied by the discharge, in both nerve and muscle, of about seventy separate electrical waves per second. Adrian and Cooper [†] have shown that interrupted stimuli at up to 230 per second can be transmitted through the spinal cord and give rise to an identical rhythm in the muscles responding reflexly to the stimulation. This evidence is in favour of the view that the stimuli causing voluntary contraction and, one must add, the tonic contractions of all skeletal muscles, are discontinuous in nature and often analogous to those which we may evoke artificially by rapid stimulation either of muscle or of its motor nerve.

[*] Gasser and Newcomer. *Amer. Journ. Physiol.*, 1921, **58**, 1.

[†] Adrian and Cooper. *Journ. Physiol.*, 1924, **59**, 61.

The problem was solved by the work of Adrian and Bronk,* by the use of electrodes so small as to pick up the electrical changes accompanying contraction in only a few fibres. The electrode was made in the form of a needle which could be stuck into the muscle of a human subject. The tracings showed the arrival of stimuli at only one or two motor units. It was found that on executing a weak contraction the muscle showed a series of regular electrical changes, each of which means the arrival, at one of the motor units, of a stimulus *via* the motor nerve fibre (Fig. 63, A). This stream of stimuli may be as slow as 6 per second, which would give little opportunity for wave summation, so we must suppose the summation in such weak stimuli to be quantal, and the smooth contraction of the whole muscle to be due to the fact that the various motor units involved are not contracting synchronously.

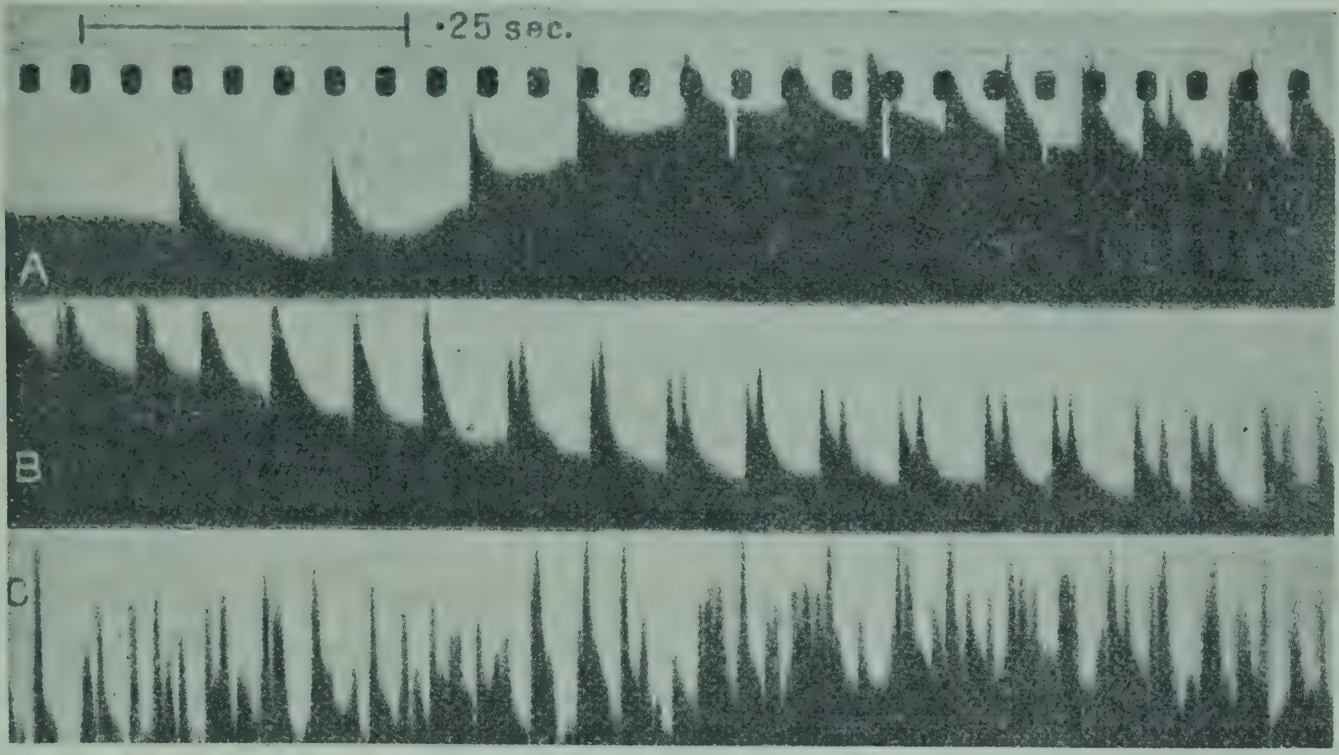


FIG. 63. Action currents in human triceps. A, weak contraction, continued at B, and showing there the entry of a new rhythm due to a fresh motor unit being excited at the same slow rate. C, powerful contraction, showing rapid rhythm of stimuli. (ADRIAN and BRONK.)

When the contraction is more powerful, fresh motor units are brought into play, and new rhythms are seen ; also the rate of stimulation is increased in each motor unit (Fig. 63, B, C), so that tetanic contractions occur. We may conclude, then, that both types of summation are concerned in voluntary contractions—in weak ones, quantal summation only, in strong ones both forms.

THE MECHANICAL ENERGY OF CONTRACTION

Whereas resting muscle is soft, flabby and presents little resistance to slight stretching, the contracted muscle is hard, rigid and resists extension. It might be thought, therefore, that the essential feature in muscular activity consisted in a change in its elastic qualities, and that contraction involved the setting up of tension which could be used for doing work or, as in the isometric twitch, degenerating into heat as the contractile state passed away. The potential energy of a muscle in a state of isometric contraction is given by the expression $Tl/8$, where l is the length of the individual fibres and T is the tension produced.† If a stimulated muscle were simply an elastic body like a spring under tension, the amount of work which it could do when allowed to shorten would be a constant amount and equal to the expression just mentioned. (We now know that owing to the " Fenn effect " it could be even greater.)

---

* ADRIAN and BRONK. *Journ. Physiol.*, 1929, **67**, 119.

† HILL. *Journ. Physiol.*, 1913, **46**, 435 ; DOI. *J. Physiol.*, 1920, **54**, 335.

It is found that under most circumstances the muscle cannot realise this maximum theoretical work, but tends more nearly to do so the slower the rate of shortening. The best method of testing the power of a muscle to do work is

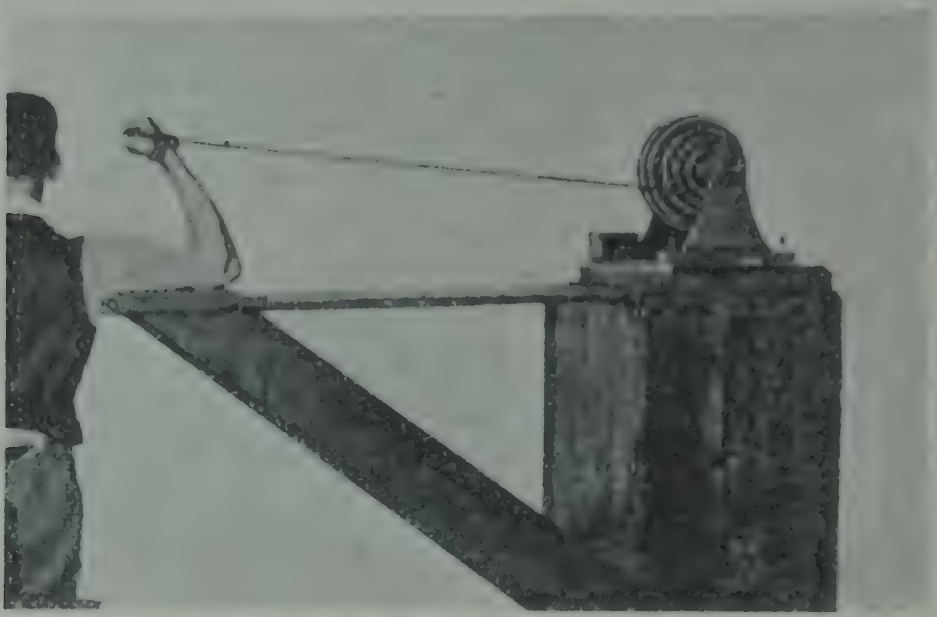


FIG. 64. Inertia flywheel for human muscle. The work done is calculated from the speed of rotation imparted to the flywheel, which weighs 35 kg., and, according to the pulley used, gives an effective mass of inertial resistance varying from 11·3 to 579 kg. (A. V. HILL. *J. Physiol.*, 1922, **56**, 19.)

to hinder its contraction by opposing it to the inertia of a mass, *e.g.* that of a heavy flywheel * (Fig. 64). If a spring under tension were made to pull against such an apparatus, the amount of work it would do would be independent of the mass moved. In muscle, however, it is found that the amount of work done against a large mass is greater than that done against a small one. The essential factor involved here is the speed with which a muscle is allowed to shorten: the greater the speed of shortening, the less, within limits, will be the work done.† This is because time is needed to liberate the energy required to move large loads. Conversely, a muscle can shorten the more quickly the smaller the

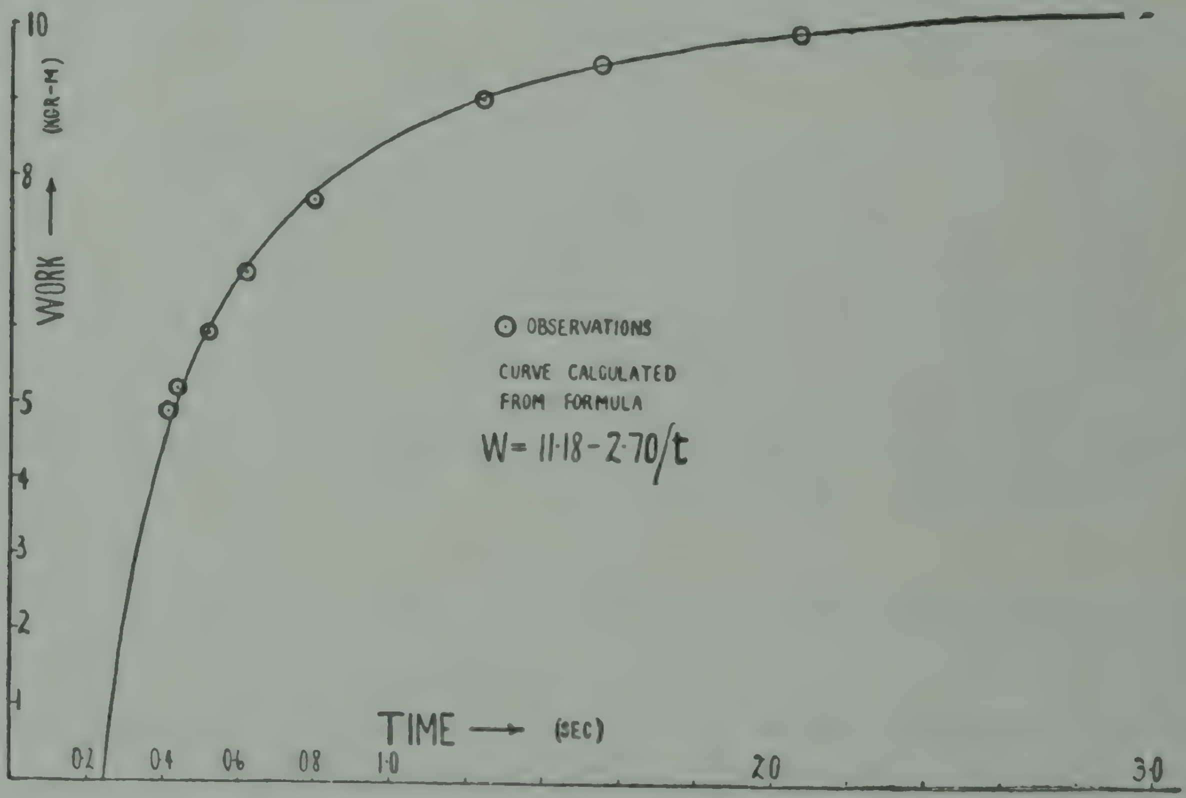


FIG. 65. Relation between maximum realisable work and duration of contraction for human arm muscles. Theoretical maximal work = 11·18 kg.-m. (A. V. HILL. *J. Physiol.*, 1922, **56**, 19.)

---

* HILL. *Journ. Physiol.*, 1922, **56**, 19.

† HILL. *Journ. Physiol.*, 1922, **56**, 19; *Proc. R.S.*, 1939, **127B**, **434**; LEVIN and WYMAN. *Proc. Roy. Soc.*, 1927, **101B**, 218.

force opposing it. Speed and force of contraction are related by the equation

$$(P + a)(V + b) = \text{constant},$$

where P = force, V = velocity, and a and b are constants; this holds for isolated frog muscle * and for human voluntary contractions.† In the case of human arm muscles the work does not approach its maximum value unless the contraction has been opposed by a mass large enough to make it occupy at least two seconds (Fig. 65). The frog's muscle is ten times quicker than that of man.

Optimal Rate of Contraction. We have seen above that the slower the speed of shortening of a muscle the greater is the amount of work done. It must be remembered however that energy is required in order to *maintain* the contracted condition in muscle.‡ There will thus be a point at which the extra energy required to maintain the contraction more than compensates for the gain in work done by increased duration of the contraction, so that for each group of muscles there will be an optimal rate at which the muscles will work with the greatest economy. We have seen that in the case of the human arm muscles the rate is ten times slower than in the isolated frog's muscle, and probably the rate will vary according to the mass of the limb involved and therefore, to a certain extent, to the size of the individual. It is a familiar experience that the effort of going upstairs becomes less as the rate is diminished, but not indefinitely, and that it may become more tiring if the rate is made too slow. The same thing applies to practically any form of exercise in which external work is being done.

FATIGUE

If a muscle be repeatedly stimulated, changes soon become apparent in the curve of contraction. The latent period is prolonged, as well as the length of the contractions; the absolute height and work done are diminished. At the same time the muscle does not return to its original length; the shortening which remains is spoken of as "*contraction remainder*." After an initial rise during the first few contractions, these diminish uniformly in height till they are no longer apparent, so that the muscle is now said to have lost its excitability. At the same time there is a great prolongation of the relaxation, so that after forty or fifty contractions several seconds may elapse before the lever returns to the base line (Fig. 66).

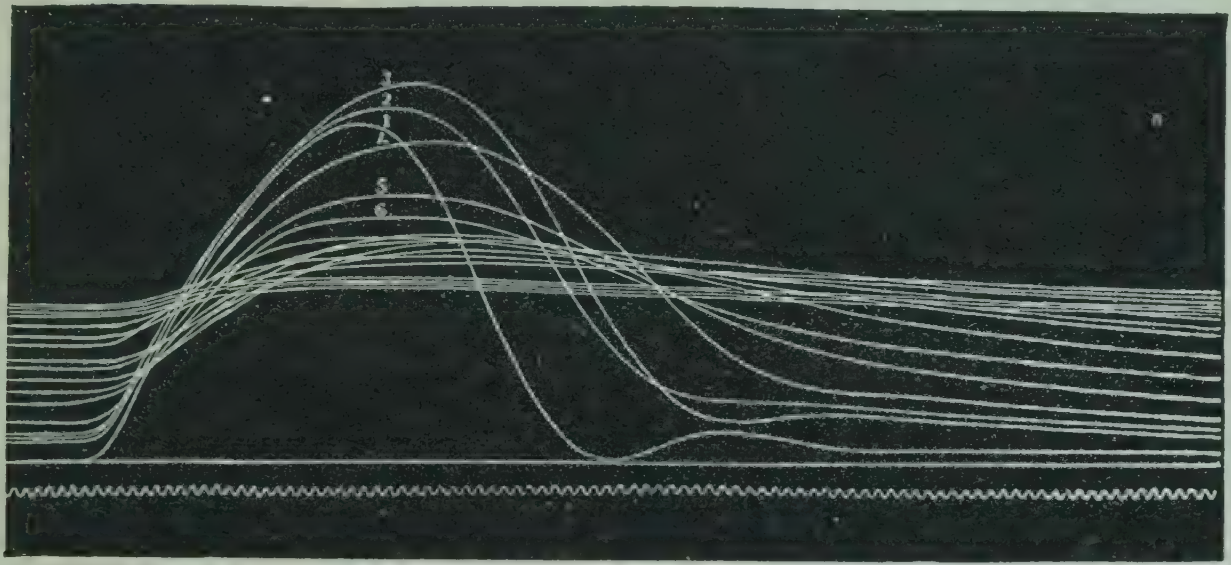


Fig. 66. Muscle Curves showing Fatigue in consequence of Repeated Stimulation. The first six contractions are numbered, and show the initial increase of the first three contractions. (Brodie.)

If left to itself, the muscle which has been exhausted by repeated stimulation will partially recover. The recovery is hastened by passing a stream of blood, or even of salt solution, through the blood vessels of the muscle. Still more important is a free supply of oxygen.

---

* A. V. Hill. *Proc. Roy. Soc.*, 1938, **126B**, 136.

† Wilkie. *J. Physiol.*, Proc., 1948, **108**, **7P**.

‡ Bronk. *Journ. Physiol.*, 1930, **69**, 306.

The phenomena of fatigue in isolated muscle depend on three factors :

(1) The consumption of the substances available for the supply of potential energy to the contractile material.

(2) The accumulation of products of the contractile process. Among these waste products is lactic acid. Fatigue may be artificially reduced by washing out the muscle with normal saline solution containing a small percentage of alkali. The beneficial effects of oxygen supply are also due in part to the removal of lactic acid and, in general, to re-establishment of the chemical *status quo.*

(3) Failure to re-establish the normal electrical polarisation of surface membranes.

In speaking of fatigue it is important to specify the tissue which is fatigued. If instead of being stimulated directly the muscle is repeatedly made to contract by stimulation of its motor nerve, fatigue sets in much sooner ; yet if the muscle be now excited directly it is seen not to be fatigued. This shows that either the nerve trunk or the motor end-plates are fatigued sooner than the muscle. Actually, the latter explanation is the correct one, since nerve trunks are practically unfatiguable.

Fatigue in the intact animal as a result of *voluntary* contractions is not strictly comparable with that of excised muscle stimulated either directly or indirectly. It may be studied in man by the use of instruments called ergographs, in which a finger, or one of the limbs, pulls up a weight or contracts against a spring, the degree of shortening or tension change being recorded. When after repeated contractions, fatigue has set in, the resulting tracing is very similar in appearance to that obtained from an isolated muscle stimulated electrically to the point of fatigue. But, in point of fact, in voluntary contraction the muscle is very far from being fatigued, as may be shown by applying electrical stimuli through the skin to it, or to its nerve, when powerful contractions result. What have been fatigued as the result of the voluntary efforts are the motor nerve cells responsible for the activation of the muscle.

We may conclude then, that the most readily fatiguable point is the motor nerve cell, next the conduction from nerve trunk to muscle, then the muscle, and last of all the nerve fibre.

CONTRACTURE

Skeletal muscle may, under certain conditions, enter into a state of prolonged reversible shortening which is distinct from tetanus, and which is called *contracture.* Thus, if a fresh frog's muscle be stimulated by repeated supramaximal opening induction shocks, the relaxation after the twitches is slow and incomplete ; it is due to a prolonged action of the stimulus.* It may in part be due to over-shortening, and passage of the fibres into the " delta state," which is only partly reversible. After many contractions also, a muscle shows another form of contracture, the " contraction remainder," due to delay in the relaxation process.

Contractures of various types can also be produced by thermal, mechanical or chemical stimuli of the most diverse description. Contractures differ from tetani and other forms of contraction in that the contractile process is not conducted through the tissue and no wave-like action potentials occur ; contractures of long duration may therefore be quite definitely localised in a muscle, the remainder of which is in a relaxed condition.

To what extent contractures may occur in the intact body is a matter of dispute ; many pathological shortenings of muscle which are seen clinically and which are sometimes called contractures, are often manifestations of tonic reflexes, fibrosis, and other phenomena not connected with real contractures. But a true prolonged contracture, called " idiomuscular contraction," is seen on mechanical stimulation of certain irritable muscles.

An interesting instance of contracture is seen in muscles after section and degeneration of their motor nerves. In such muscles, stimulation of the *vaso-dilator* nerve (or of the sensory nerve) to the muscle causes a prolonged contracture. This, which is sometimes called the Vulpian-Heidenhain-Sherrington phenomenon, is of great theoretical interest. It was discovered by Vulpian and Philippeaux in 1863 that, some days after section of the hypoglossal nerve (motor to the tongue), stimulation of the chorda tympani fibres in the lingual nerve (vaso-dilator to the tongue) caused contracture of the tongue muscles. In the normal state, stimulation of the chorda tympani produced no effect at all on the tongue muscles. It was shown by Sherrington * that the phenomenon is obtainable under comparable conditions from any muscle

---

* BREMER. *J. Physiol.*, 1932, **76**, 65.

The key to this apparent paradox is to be found in the theory that the action of vaso-dilator nerves is due, at least in part, to the liberation at their terminations in the tissues, of acetyl choline.† The muscle when deprived of its motor nerve supply is thrown into contracture by acetyl choline, as may easily be demonstrated, whereas the normally innervated muscle is less sensitive to its action. By injection of acetyl choline in large doses (*e.g.* 1–2μg.) direct into the arterial blood stream to a normal muscle, however, a quick contraction, followed by a slow contracture, results.‡ The effect of acetyl choline on the denervated muscle, in accordance with this hypothesis, is enhanced by eserine and abolished by atropine. What happens under the conditions of the experiment, therefore, is that when the vaso-dilator nerve is stimulated, acetyl choline is liberated in the region to which the nerve is distributed, *i.e.* in the muscle. This substance causes dilatation of the bloodvessels there, both in the normal muscle and in the denervated muscle. But in the denervated muscle the acetyl choline also produces the effect of contracture, which it does not do in the less sensitive normal muscle.

For Reference

Gasser (1930). Contractures of Skeletal Muscle. *Physiol. Rev.*, **10**, 35.

Rigor Mortis. All muscles after removal from the body, or if left in the body after general death, lose their excitability after a time, and this loss is succeeded by the phenomenon known as *rigor mortis*. The muscle, which was previously flaccid, shortens, though not very powerfully. Whereas the living muscle is translucent, supple and extensible, it becomes in the process of rigor opaque and much less extensible, the modulus of elasticity, measured along the fibres, increasing about tenfold.§ The rate of onset of rigor is greatly accelerated by treating the muscles with solutions of caffeine, arsenates or chloroform vapour, but is not much affected by temperature.§ When rigor has been established, the reaction of the muscle is *usually* found to have changed from a slightly alkaline (pH 7·3) to a distinctly acid one (pH 5·8), due to the presence of lactic acid. But acidification is not a necessary condition for rigor to occur. From this condition of rigor there is no recovery of contractility, though the muscle softens before putrefaction sets in. There can be no doubt that the change in consistence of the muscle, and probably also its shortening in rigor, are due to changes in the muscle proteins. Both conditions can be imitated by heating the muscle. After severe muscular fatigue, as in hunted animals, rigidity may come on almost immediately after death.

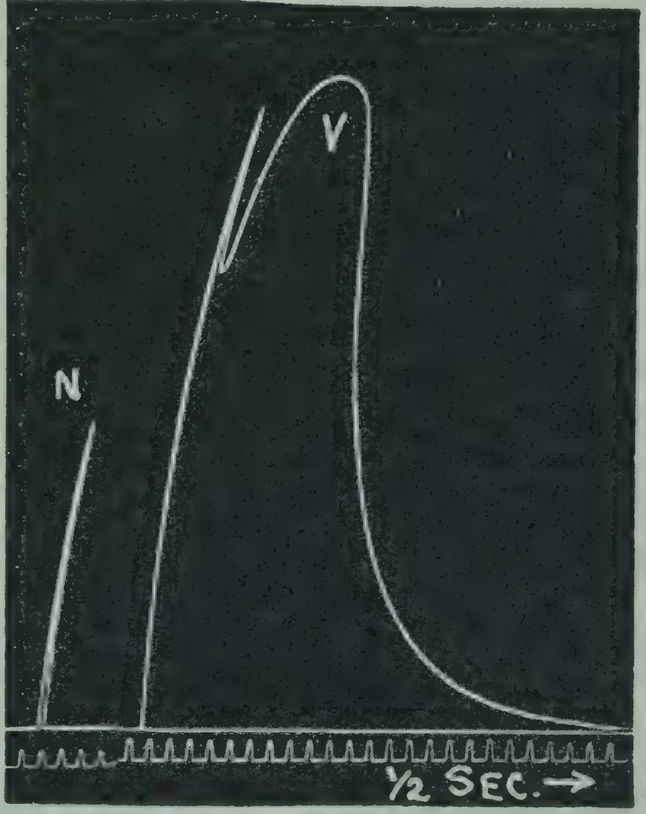


Fig. 67. Contractions of frog's sartorius. N, before, and V, twenty minutes after, soaking in Ringer's solution containing 1 : 100,000 veratrine. The records were taken on a slowly moving surface, as the time tracing shows.

THE ACTION OF CHEMICAL SUBSTANCES

Salts. The membrane of muscle fibres is permeable to the $K^{\cdot}$ ion, but impermeable to $Na^{\cdot}$ or larger cations ; it is also permeable to small anions, such as Cl', HCO_3', etc. The question of the distribution of ions inside and outside the fibre is too complex to discuss here.||

---

* Sherrington. *J. Physiol.*, 1894, **17**, 252.
† Dale and Gaddum. *J. Physiol.*, 1930, **70**, 109.
‡ Brown. *J. Physiol.*, 1937, **89**, 438.
§ Smith. *J. Physiol.*, 1939, **96**, 176.
|| Boyle and Conway. *J. Physiol.*, 1941, **100**, 1.

It has been supposed that potassium accumulates inside the fibres t replace sodium, which is actively expelled, and is not able to return. The action of sodium salts on muscle is of considerable interest. A 0·65 per cent NaCl solution has an excitatory effect upon the muscle, so that single induction shocks may cause contracture. The same excitatory effect is still better marked with solutions of sodium carbonate or oxalate which reduce the free $Ca^{\cdot\cdot}$ of the muscle. If a thin muscle, such as a frog's sartorius, be immersed in a solution containing 0·5 per cent. NaCl, 0·2 per cent. Na_2HPO_4, and 0·04 per cent. Na_2CO_3 (Biedermann's fluid), the muscle enters into a series of frequent contractions. This excitatory action of sodium salts is neutralised by the addition of traces of calcium salts.

Potassium salts if injected direct into the artery supplying a muscle, in doses of a few mg. cause a brief twitch.[*] Yet if a muscle is soaked in a solution of KCl, it quickly loses its excitability. If a muscle be transfused with suitable fluids containing small amounts of potassium salts, it shows all the signs of fatigue, which may be removed by washing out the potassium salts.

If a frog's muscle be isolated and kept moist for a few hours, it often shows a complete loss of excitability, but regains it after a short immersion in Ringer's fluid, and does not become non-irritable again during its survival. The loss of excitability is apparently due to the passage of potassium salts from the inside to the outside of the fibre.[†]

OTHER DRUGS. Of the drugs that have a direct action on muscle, the most remarkable is veratrine. If a muscle is soaked in a 1 : 100,000 solution of this alkaloid, a single stimulus results, not in a short twitch, but in a sustained "contracture," which may last for many seconds (Fig. 67). The effect appears to be due to an increase in the duration of the excitatory state or to scattered asynchronous tetani of single fibres.[‡] For some time after the "contracture" has died down the effect of stimulation is to give an ordinary twitch. Formaldehyde, barium salts, and many other substances produce similar contractures.[§]

Quinine increases the response to a single maximal stimulus, but also increases the refractory period, so that response to a tetanus is diminished. Eserine and veratrine do not produce their effects after quinine.[||]

Chloroform and caffeine cause the muscle to shorten and quickly to pass into *rigor mortis*. Barium salts have a similar though less marked effect.

CHEMICAL COMPOSITION OF SKELETAL MUSCLE

Muscle contains about 75 per cent. of water. The chief solid constituents are the proteins (20 per cent.) which form the bulk of the cellular structure of the tissue. Of the remaining solids, several constituents referred to later are of great importance, though only present in small amounts.

MUSCLE PROTEINS. The contents of the living fibres, once thought to be fluid, are now recognised to be in the state of a weak gel; this is shown by their behaviour when expressed by micro-dissection operations, and also by a consideration of certain physical properties of the muscle. By mincing the living muscle with ice and a little salt and subsequently filtering under pressure, a viscous fluid formerly called "muscle plasma" can be obtained. This is not, as once thought, the unaltered cell contents, but a partial solution of the gel in dilute salt solution. If the temperature of the extract rises, it clots to a firm gel, but the nature of the coagulation is unknown.

The older investigations of Halliburton and others have been extended,[¶] [**] and it appears that the living muscle contains five proteins, called respectively *myogen*, *myosin*, *globulin X*, *myoalbumin* and *stroma proteins*.

Myogen, a mixture of albumins, is obtained by dialysing the "plasma," which precipitates the myosin, and subsequently filtering. It is soluble in water and has an isoelectric point of pH 6·3. It is readily denatured at the isoelectric point. It represents about 20 per cent. of the total proteins. Part of it consists of enzymes.

---

[*] BROWN. *J. Physiol.*, 1937, **91**, 4.P.

[†] HORTON. *J. Physiol.*, 1930, **70**, 389.

[‡] KRAYER and ACHESON. *Physiol. Rev.*, 1946, **26**, 383.

[§] GASSER. *Physiol. Rev.*, 1930, **10**, 35.

[||] HARVEY. *J. Physiol.*, 1939, **95**, 45.

[¶] SMITH. *Proc. Roy. Soc.*, 1930, **105B**, 597; 1937, **124B**, 136.

[**] WEBER. *Ergebn. d. Physiol.*, 1934, **36**, 109.

Myosin, a globulin, is obtained by precipitation from a slightly alkaline (pH 7·2) saline extract by neutralisation and dilution. The precipitate is redissolved in dilute salt solution at pH 7·0.* Its solutions are extremely viscous and it readily forms a gel when the salt concentration is low. Its isoelectric point is dependent on the salt concentration, but is about pH 5·1. The globulin is insoluble in water, and, between pH 5 and 6, also insoluble in salt solutions. It is completely precipitated in 0·04 m. KCl at pH 7·0. About 40 per cent. of the total muscle protein is myosin, and this is largely present in the Q-discs. Here, owing to the rod-like nature of its molecules, it confers on the muscle its elasticity, its thermo-elastic properties and the known relationship between damped and undamped elasticity. According to Szent-Gyorgyi † myosin can be fractionated into two proteins, *myosin* and "*actin*," which combine to form a complex called "*actomyosin*." This has different properties from its components, and in particular can be caused, when in gel form, to "contract" when brought into contact with A.T.P. or K ions. The shortening is, however, accompanied by volume shrinkage or synæresis of the gel, and is not strictly comparable with the shortening of the contractile material in muscle.

Globulin X, only slightly soluble, is completely precipitated in 0·005 m. KCl at pH 7·0. It represents about 20 per cent. of total protein.

Stroma proteins are insoluble even in concentrated salt solutions, but it is uncertain how much, if any, of these represent constituents of the living fibre.

It is interesting to note that the amount of protein capable of extraction from muscle by means of salt solution is greater (90 per cent.) when the muscle is quite fresh. When rigor has set in only 30 per cent. of the total protein is thus extractable and ‡ it would seem that the processes of fatigue and rigor, like that of the coagulation of fresh extracts, is the result of a process of denaturation which takes place in the protein before or after death. The hydrogen ion concentration of fresh resting muscle is probably about pH 7·0 § so that the muscle proteins would be on the alkaline side of their isoelectric point, *i.e.* they would be present as salts of bases and be ionised as anions. Although coagulation can be produced in a fresh muscle by immersion in acid, it appears to be different from true rigor mortis, which, moreover, is not necessarily associated with an acidification of the muscle.

OTHER CONSTITUENTS OF MUSCLE. Of the remaining constituents of fresh resting muscle the most important are glycogen, lactic acid, and three types of phosphorus compounds, viz. creatine phosphate, hexose phosphates and the adenosine phosphates ; all of them appear to participate in the enzymic chemical changes which are associated with contraction.

Glycogen is present up to 0·5 to 1 per cent. ; embryonic muscle contains more.

Lactic acid ($CH_3.CHOH.COOH$) is the *l*-(+) isomer. It is only present in amounts of the order of 0·02 per cent. when muscle is fresh, well oxygenated and resting. When the muscle is injured, and under various other conditions, glycogen is rapidly converted into lactic acid.

The *hexose phosphates* are discussed on p. 34, and *creatine phosphate, adenosine diphosphate* (A.D.P.) and *adenosine triphosphate* (A.T.P.), on pp. 52 and 68.

Creatine phosphate is very readily broken down by acid, or by enzymes on injury of the muscle, into creatine and phosphoric acid. It has been synthesised.‖ In invertebrates it is replaced by arginine phosphoric acid, which has similar functions. Phosphocreatine is a stronger acid than phosphoric, and so is an important base-binder, the base being set free when it hydrolyses.

The phosphoric acid compounds are all important in connection with the chemical changes that occur in muscle in relation to contraction. Indeed, since phosphorylation plays a dominant *rôle* in many processes of intermediary metabolism, the general reaction

$$\text{A.T.P.} + H_2PO_3\text{-acceptor} \rightleftharpoons \text{A.D.P.} + H_2PO_3\text{-donator}$$

by which H_2PO_3 is moved into and out of molecules (by the action of appropriate enzymes), is of the greatest biochemical significance. Moreover, as already explained (p. 52) some of the phosphate groups carry a large amount of energy which can either be set free to do work, or transferred withou diminution to other compounds.

---

* EDSALL. *J. Biol. Chem.*, 1930, **89**, 289.

† SZENT-GYORGYI (1947). Chemistry of Muscular Contraction. (Acad. Press Inc., N.Y.).

‡ SAXL. *Beitr. chem. Physiol. u. Path.*, 1907, **9**, 1.

§ FURUSAWA and KERRIDGE. *Journ. Physiol.*, 1927, **63**, 33.

‖ ZEILE and FAWAZ. *Ztschr. f. Physiol. Chem.*, 1938, **256**, 193.

Porphyrin Derivatives. The red muscles contain an iron-holding pigment which closel resembles hæmoglobin; it has been isolated in crystalline form* and called *myoglobin*, o myo-hæmoglobin. It probably acts as a temporary oxygen store for the muscle. In addition all skeletal muscle contains the Fe-porphyrin pigments *cytochromes* a, b and c † and also cytochrome oxidase and dehydrogenase, all of which have important functions in oxidation.

Other constituents are :—*Hypoxanthine*, which is present in muscle extracts, is chiefly derived from the adenylic acid of the tissue, and only to a small extent from the nuclein of the cell nuclei.

Carnosine is a dipeptide, β-alanyl histidine. It is present to about 0·3 per cent., but its functions are unknown. In some animals it is replaced by a methyl derivative called anserine.

Inositol ($C_6H_{12}O_6 . 2H_2O$) is hexahydroxybenzene; its significance is unknown.

Fat, Lecithin and Cholesterol in traces.

Inorganic Constituents. The ash of muscle consists chiefly of potassium phosphate (K, = about 0·3 per cent. of the wet weight), with traces of Ca, Na, Mg and Fe in the fresh tissues. The high K content helps to keep the myosin in a suitable physical state, perhaps in solution: it also plays an important part in the maintenance of electrical polarisation of the plasma membrane. Most of the bases are normally combined with the proteins, while the phosphate is present in the various organic forms mentioned above. We therefore find that the phosphate and potassium do not diffuse out from the living resting muscle, though they do so to some extent after stimulation. This, as we shall see later, is because of temporary liberation from organic combination.

Enzymes. Various enzymes (some crystalline) have been prepared from muscle, and play an important part in bringing about the chemical changes which occur in connection with contraction; thus there are enzymes which produce and hydrolyse the phosphoric esters, and by the successive action of many enzymes we may get conversion of glycogen into lactic acid, or final oxidation to CO_2 and water.

THE CHEMICAL CHANGES WHICH ACCOMPANY CONTRACTION

The energy of the contraction of muscle must ultimately be derived from chemical changes occurring in it. In seeking out the nature of these changes four methods are open to us:

(1) We can examine the changes in the muscle either *in situ*, or surviving.

(2) We can investigate the changes in the blood or in an artificial medium surrounding the muscle. Muscle may be exposed in nitrogen or oxygen in a confined space, and its gaseous interchanges during rest and activity compared. The muscle may be suspended in Ringer's solution, with or without oxygenation, or we may determine the change in the composition of the blood in passing through the muscle.

(3) In an attempt to analyse the net changes shown by (1) and (2), the chemical reactions brought about by the action, on given substrates, of minced muscle, extracts, or reconstituted enzyme systems can be followed.

(4) We may compare the changes in the total metabolism of the animal during rest and muscular work as a clue to the muscular metabolism itself.

We may begin by dealing with the evidence afforded by these methods of investigation and then see how far our conclusions are borne out by experiments on the metabolism of the whole animal. The results of method (3) will be discussed more fully later on.

It must be admitted at the outset that, despite the enormous labours of some of the world's most eminent biochemists, there are still many gaps in our knowledge. The outstanding features of the net chemical changes after contraction are :—

(i) Oxygen is used, and carbon dioxide produced.

(ii) Glycogen is reduced in amount.

Further inspection suggested that these changes divide into two stages :—

(i) Glycogen is reduced and, in absence of oxygen, lactic acid is formed.

(ii) Lactic acid disappears, O_2 is used up and CO_2 is formed.

---

* THEORELL. *Bioch. Ztschr.*, 1932, **252**, 1.

† KEILIN. *Proc. Roy. Soc.*, 1925, **98B**, 312.

Much experimental work has been concerned with changes that occur independently of oxygen. Under anaerobic conditions it is beyond question that the main changes consist in an increase of lactic acid, with a concomitant increase in hexosemonophosphate and inorganic phosphate, and a reduction in creatine phosphate. These are probably intermediary changes, and are replaced or succeeded in the normal muscle by restitutive reactions in which oxygen is involved; at all events, when oxygen is present these changes are diminished or suppressed.

The changes by which glycogen is built up, and then ultimately disappears by conversion into CO_2 and water are probably much the same in all tissues, and represent the essence of our knowledge of intermediary metabolism of carbohydrate. The focal point of these changes is represented by pyruvic acid, though a more evident intermediary is lactic acid. We may regard lactic acid as the end-product of glycogen breakdown when oxygen is not present, *i.e.* under anaerobic conditions, and we will consider these conditions first.

THE LACTIC ACID CYCLE IN SURVIVING MUSCLE

FORMATION OF LACTIC ACID. The lactic acid content of muscle is estimated by freezing in liquid air, and mincing in an ice-cold solution of trichloracetic acid. After removal of carbohydrate from the extract the lactic acid of the filtrate is determined.

It is easy, by modern methods, to confirm the fact, originally elicited by Fletcher and Hopkins,* that fresh resting muscle contains only minimal amounts of lactic acid, provided that care is taken to avoid injury and to keep its temperature low until it has been killed by the reagent; the rate of formation of lactic acid is quite slow provided the muscle is uninjured and at rest. If the muscle has been stimulated and made to contract immediately before being killed in this way, its lactic acid content is found to be raised considerably, and a few seconds *after* stimulation it is higher still. The total rise in mammalian muscle may be from a resting value of 0·02 per cent. up to a final value of 0·5 per cent. These experiments prove that lactic acid is formed when a muscle contracts, and continues to be formed for a short time after the contraction is over. It is also easy to show that as the lactic acid increases the glycogen content of the muscle proportionally diminishes. If the muscle be damaged, a steady formation of lactic acid takes place, which is more rapid in the first few hours after death, but continues until the muscle passes into rigor.

With the completion of rigor, frog's muscles are found to contain about 0·4 per cent. lactic acid. Above this the amount does not increase. The onset of rigor, and the rate of production of lactic acid, are quickened if the muscle be kept warm and if oxygen be excluded. It is interesting to note that the amount of lactic acid found in rigid muscle is about the same, whatever the previous history of the muscle. Thus, if the fresh muscle be finely minced it is found to contain about 0·2 per cent. lactic acid. If, however, it be allowed to stand after mincing, there is a slow production of lactic acid up to the maximum 0·4 per cent. Again, a muscle which has been tetanised to exhaustion contains about 0·2 per cent. lactic acid. When allowed to undergo rigor, the amount rises to about 0·4 per cent. This maximum depends on the fact that the accumulation of lactic acid inhibits its further formation. If care be taken to neutralise the acid as it is formed by immersing the muscles in weak alkaline solutions, especially of phosphates, *e.g.* Na_2HPO_4, and to exclude oxygen, the whole of the glycogen disappears and as much as 1·2 per cent. lactic acid can be obtained.

In life, the lactic acid which is formed in muscles by their contraction is immediately neutralised. The process is complex, and so perfect that little if any change of reaction occurs until fatigue begins to set in.

In this neutralisation the inorganic salts of the muscle take but a small part, but an important *rôle* in the process is played by the hydrolysis of creatine phosphate and by the buffering action of the proteins of the muscle, which exist as ionised compounds of protein and potassium or sodium—as in the following equation :

---

* FLETCHER and HOPKINS. *J. Physiol.*, 1907, **35**, 247; *Proc. Roy. Soc.*, 1917, **89B**, 444.

$$\underset{\text{(sodium proteinate)}}{Na^+ + P^-} + \underset{\text{(lactic acid)}}{H^+ + L^-} \rightarrow \underset{\text{(sodium lactate)}}{Na^+ + L^-} + \underset{\text{(undissociated protein)}}{HP}$$

Two points regarding lactic acid formation are important. First, the lactic acid formation does not precede or accompany a contraction but *follows* it. Hence it cannot be the cause of the contraction. Second, lactic acid formation occurs independently of any oxygen utilisation and is, indeed an anaerobic phenomenon.

Source of the Lactic Acid. The disappearance of glycogen as lactic acid appears, points to that substance as the precursor of the lactic acid. The conversion of glycogen (or glucose) to lactic acid is called *glycolysis.* Moreover, the maximal amount of lactic acid obtained is never greater than corresponds to the original glycogen content. It might be expected that the glycogen would first be converted to glucose, but this is not what happens : the enzymes of muscle form lactic acid *more rapidly from glycogen than from glucose.* The change requires the presence of phosphates, and takes place by a series of complex phosphorylated intermediary stages, in which hexose and triose phosphates (pp. 31, 34) are important intermediaries.

The following scheme represents in an abridged form some of the principal changes involved, and will be called "*the lactic acid cycle.*" It is held that in the living muscle all these steps are reversible.

Glycogen
⥮
Hexose monophosphates
⥮
Hexose diphosphate
⥮
Triose phosphate
⥮
Pyruvic acid
⥮
Lactic acid.

CHEMICAL CHANGES UNDER AEROBIC CONDITIONS

The arrangements for the supply of oxygen to muscles in the living body are extremely good, and when a muscle carries out only moderately powerful and moderately frequent contractions under these conditions, little if any lactic acid appears. It was at one time supposed that this was because the lactic acid was formed, as under anaerobic conditions, but that in presence of a full oxygen supply this was in part ($\frac{1}{5}$) oxidised and in part ($\frac{4}{5}$) resynthesised to glycogen.

It is now considered that the cross-roads of carbohydrate metabolism is not lactic acid, but pyruvic acid. In absence of oxygen, pyruvic acid is reduced to lactic acid by lactic dehydrogenase, so that under anaerobic conditions, lactic acid accumulates ; in presence of oxygen, pyruvic acid is, instead, completely oxidised to CO_2 and water, the net reaction being :—

$$CH_3CO.COOH + 50 = 3CO_2 + 2H_2O$$

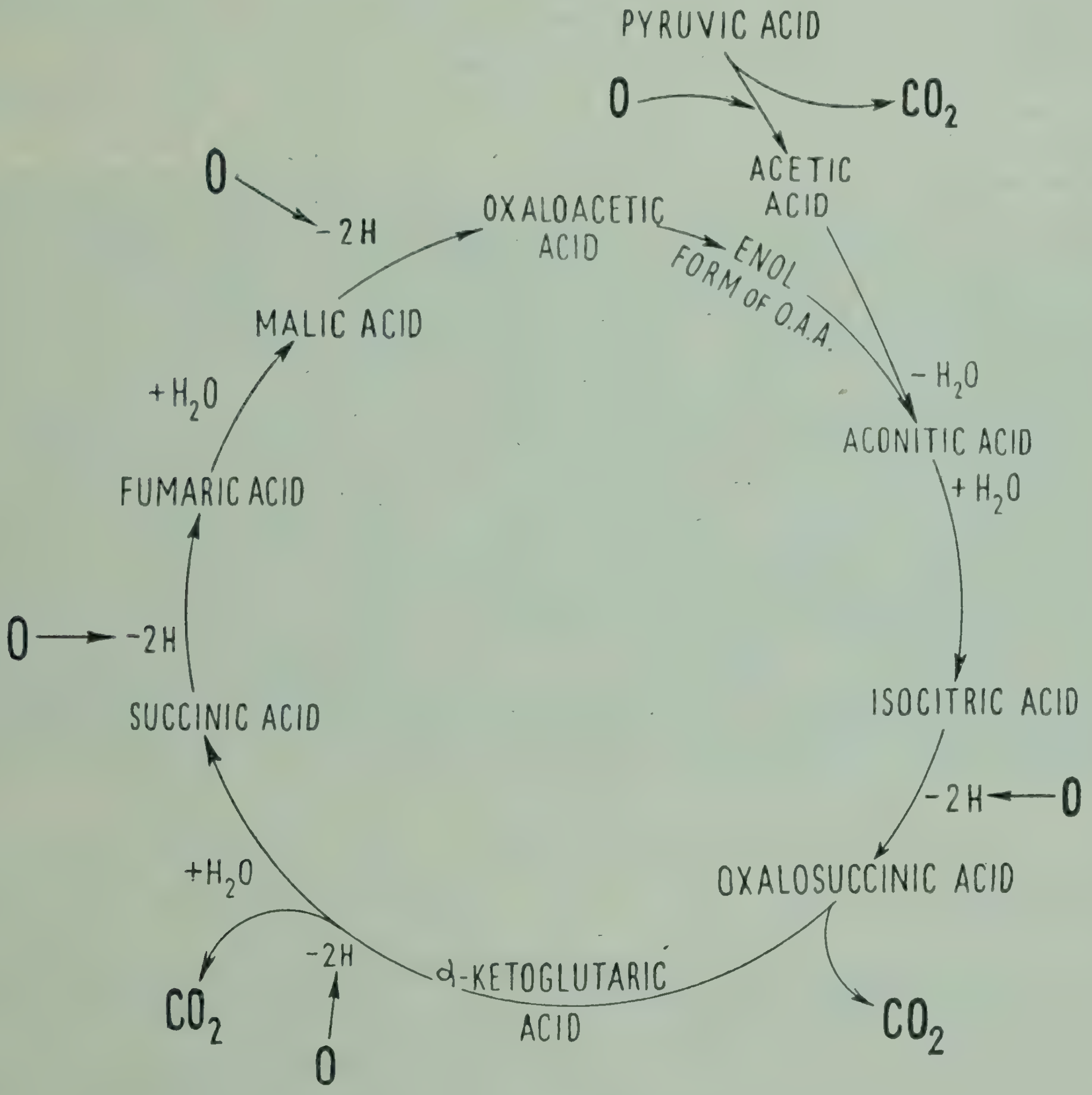
This final result, which represents the end stage of the oxidation of carbohydrate to CO_2 and water, is reached by an interesting series of intermediate reactions, each controlled by a specific enzyme. The details of these reactions will be found in text-books of biochemistry and can only be briefly summarised here.

The first change is the oxidative decarboxylation of pyruvic acid with immediate liberation of the first molecule of CO_2 by the reaction :—

$$CH_3.CO.COOH + O = CH_3COOH + CO_2$$

For this oxidative decarboxylation to occur, coenzyme I, A.D.P., phosphate, Mg ions and the pyrophosphate of vitamin B_1 are all essential. The last named is sometimes called co-carboxylase, and in conditions of lack of vitamin B_1, pyruvic and lactic acids accumulate in the brain, because the pyruvic acid cannot be removed by oxidative decarboxylation.

The next step in the final oxidation of pyruvic acid is that the acetic acid formed as above enters the so-called "tricarboxylic acid cycle." This is a cyclic series of changes which begins with the 4- carbon *di*carboxylic acid, oxaloacetic acid. The acetic acid, by condensing with the *enol*- form of oxaloacetic acid, forms the 5- carbon *tri*-carboxylic acid aconitic acid; this then passes into isocitric acid by adding on H_2O; then by dehydrogenation this acid is converted into oxalosuccinic acid; this is then decarboxylated to give the 5- carbon *di*carboxylic, α-ketoglutaric acid, together with a second molecule of CO_2; the α-ketoglutaric acid is then further decarboxylated, giving the third molecule of CO_2 and succinic acid (a 4- carbon dicarboxylic acid): the succinic then passes by dehydrogenation to fumaric acid, this by hydrolysis to malic acid, which then by dehydrogenation gives oxaloacetic acid with which the cycle began. A small amount of oxaloacetic acid can thus, by repeating this cycle of changes over and over again, continue to oxidise pyruvic acid to CO_2 and water, one molecule of pyruvic acid being taken up at each cycle and yielding 3 molecules of CO_2 as shown in the scheme below:—



The structural formulæ of the intermediary compounds in the cycle are given on p. 31. Each of the changes is effected by a specific enzyme, and in the 5 cases where these are dehydrogenases, the 2H abstracted is passed on in each case through the cytochrome oxidase system, as explained on p. 100, to an atom of oxygen. Hence, as shown in the scheme, we get 3 molecules of CO_2 formed and 5 atoms of oxygen used for each molecule of pyruvic acid broken down, which summarises in the equation:—

$$CH_3CO.COOH + 5O = 3CO_2 + 2H_2O.$$

In the complete oxidation of a carbohydrate, *e.g.*, glucose, the volume of oxygen used is always equal to that of the CO_2 produced, but in the above equation an atom less of oxygen is required. The explanation is that, when half a molecule of a hexose is being broken down under aerobic conditions, there is, in the stages leading to pyruvic acid, an oxidation of 2H (*via* cytochrome) by O, as a consequence of dehydrogenation reactions: hence, in all $3O_2$ is used for oxidation of half the hexose molecule, and $3CO_2$ produced, as explained on p. 145.

Energy Changes in Muscle

The energy of muscular contraction is ultimately derived from the complete oxidation of glycogen to CO_2 and water. It is true that, in absence of oxygen, muscle can for a while develop contractile energy, but this is done (1) by drawing upon stores of energy which are laid down in the form of energy-rich phosphate bonds (*v.* p. 52), in the form of A.T.P. and creatine phosphate, and (2) by energy set free by glycolysis, by which these phosphate bonds are restored. This last source of energy, however, only results in a loss of free energy of about 57 cal. per unit mole of hexose derived from glycogen, of which about 30 cal. is stored as energy-rich phosphate bonds, *i.e.* 3 (~). In complete oxidation of the 6-carbon unit of glycogen, however, there is a loss of free energy about twelve times as much as in glycolysis, *i.e.* of about 720 cal. which would possibly be equivalent to creation of about 36 (~).

We do not know at exactly what stages in the oxidative breakdown of pyruvic acid the new energy-rich bonds are generated, but it is certain that in the long run it is the oxidative changes that yield most of the energy, and this appears partly as heat and partly in the storage form of energy-rich bonds, by the utilisation of which future contractions are provided for.

Probable Sequence of Reactions. A sequence of reactions of glycolysis of glycogen is given in detail in the table below, and the enzymes concerned are shown.

Reaction No.	Enzyme system.	Reacting systems
1	Myosin I + $Ca^{\cdot\cdot}$	A.T P. → A.D P. + H_3PO_4
2	Phosphorylase (protein-Mg-adenylic acid).	Glycogen + H_3PO_4 ⇌ glucose 1-phosphate.
3	Phosphoglucomutase (protein-Mg).	Glucose-1-phosphate ⇌ glucose 6-phosphate.
4	Oxoisomerase (protein-M*).	Glucose 6-phosphate ⇌ fructose 6-phosphate.
5	Phosphohexokinase (protein-ATP)	Fructose 6-phosphate + A T.P. ⇌ fructose 1:6-phosphate + A.D P.
6	Adenyltriphosphatase.	A.T.P.+creatine ⇌ A.D.P.+creatine phosphate.
7	Zymohexase (protein-M*)	Fructose 1:6-phosphate ⇌ dioxyacetone phosphate. Fructose 1:6-phosphate ⇌ 3-glyceraldehyde phosphate.
8	Phosphotriose isomerase (protein-M*).	Dioxyacetone phosphate ⇌ 3-glyceraldehyde phosphate.
9	Nonenzymatic.	3-Glyceraldehyde phosphate + H_3PO_4 ⇌ 1:3-diphosphoglyceraldehyde.
10	Triosephosphate dehydrogenase (protein-DPN) (=Co.I.)	1:3-Diphosphoglyceraldehyde + Co.I. ⇌ 1:3-diphosphoglycerate + Co.I.H_2.
11	A phosphokinase.	1:3-Diphosphoglycerate + A.D.P. ⇌ 3-phosphoglycerate + A.T.P.
12	Phosphoglycere mutase (protein-M*).	3-Phosphyglycerate ⇌ 2-phosphoglycerate.
13	Enolase (protein-Mg).	2-Phosphoglycerate+H_2O⇌ phosphopyruvate.
14	Phosphopyruvate phosphatase (protein-ADP-Mg).	Phosphopyruvate + A.D.P. → pyruvate + A.T.P.
15	Lactic Dehydrogenase.	Pyruvate + Co.I.H_2 ⇌ Lactate + Co.I.

* M = metal.

It will be seen from these reactions that phosphoric acid from A.T.P. is passed on from one compound to another, and ultimately restored.

In formation of lactic acid from glucose (glycolysis), the first enzyme is *hexokinase.* This causes the reaction : Glucose + A.T.P. ⇌ glucose 6-phosphate + ADP. Thereafter the reaction proceeds as in No. 4 onwards. It should be noted that the reduced Co.I formed in reaction 10 is oxidised again in reaction 15, when anaerobic conditions prevail, the 2H lost being used to reduce pyruvic to lactic acid. When oxygen is present, reaction 15 does not occur, but instead we get the two reactions :—

(*a*) $CH_3.CO.COOH + 5O = 3CO_2 + 2H_2O$, as explained above.
(*b*) $Co.I.H_2 + O = Co.I + H_2O$,

the net result of which is that $3O_2$ is used and $3CO_2$ formed.

The reverse change of conversion of lactic acid into glycogen is slow, and does not occur at all in the absence of oxygen. When oxygen is available, and if the muscle is allowed to rest, the lactic acid is removed, and glycogen returns ; it is, however, so slow a process that we may doubt whether such recovery plays a large part in restoring glycogen to muscles *in vivo.*

The Oxygen Usage of Muscle. If we stimulate a muscle repeatedly in an atmosphere of oxygen, it ultimately becomes fatigued and incapable of further contraction. On allowing the muscle to rest, however, it recovers considerably and can execute a further series of contractions, again resulting in fatigue. The recovery is hastened by washing the muscle in oxygenated Ringer's solution.

The muscle can also contract equally well to start with if placed in an atmosphere of nitrogen. Fatigue finally sets in ; but on cessation of the stimulation the muscle in nitrogen does not recover, or recovers very little, and soon passes into rigor * (Fig. 68).

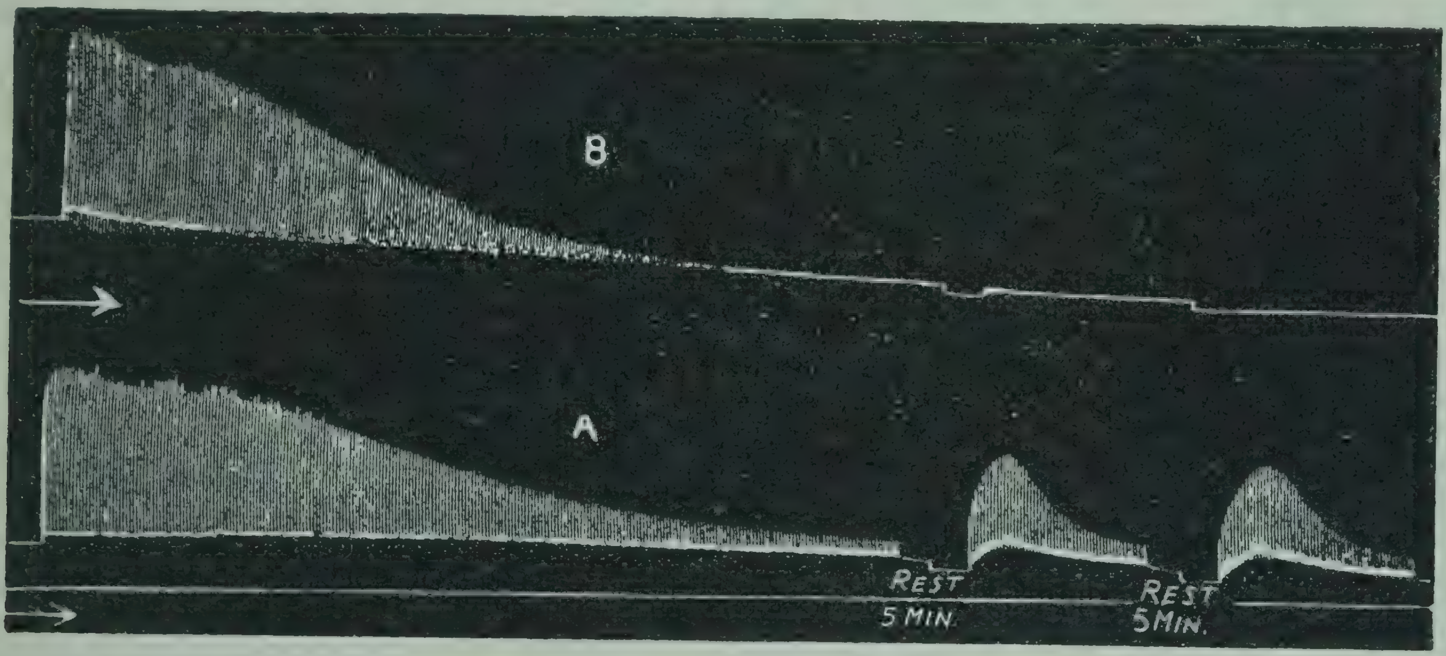


FIG. 68. Fatigue in a Pair of Sartorii. (FLETCHER.)
Max. break shocks one per second. Load 6 grammes. Temp. 19° C.
A. Exposed to oxygen. *B.* Exposed to nitrogen.

These facts suggest that, although oxygen is not essential for contraction it is necessary for recovery. It is therefore of interest to examine the oxygen usage of excised frog's muscle, both resting and fatigued.

For the investigation of the oxygen utilisation of isolated tissues much use has been made of the *microrespirometer,*† the principle of which is illustrated by Fig. 69. The two small flasks A and B are of equal size, and are connected each with one limb of a fine-bore manometer filled with kerosene and provided with a millimetre scale. On the bottom of each of the flasks is a receptacle (*a*, *b*) in which are placed a few drops of potash solution ; the tissue, on a thin sheet, or minced, is floated in Ringer's or other suitable solution in the bottom of one flask, the other containing Ringer's solution only. The bottles are placed in a thermostat and filled with oxygen. The menisci of the manometer are read, and the whole apparatus is continuously agitated in a horizontal direction to and fro to keep the solution saturated with oxygen. At intervals the reading of the meniscus in each limb of the manometer is taken. The level of the kerosene ascends in the manometer on the side connected with the bottle containing the tissue, and falls on the other side. The carbon dioxide produced by the tissue is absorbed by the potash, so that the movement of the kerosene is entirely due to the oxygen absorbed, the volume of which can thus be calculated. Thus the oxygen utilisation of the tissue under given conditions can be found over any required period of time.

---

* FLETCHER. *J. Physiol.*, 1902, **28**, 474.

† KROGH. *Bioch. Ztschr.*, 1914, **62**, 266. There are many patterns of microrespirometer. *v.* DIXON. *Manometric Methods*, 1934. (Cambr. Univ. Press.)

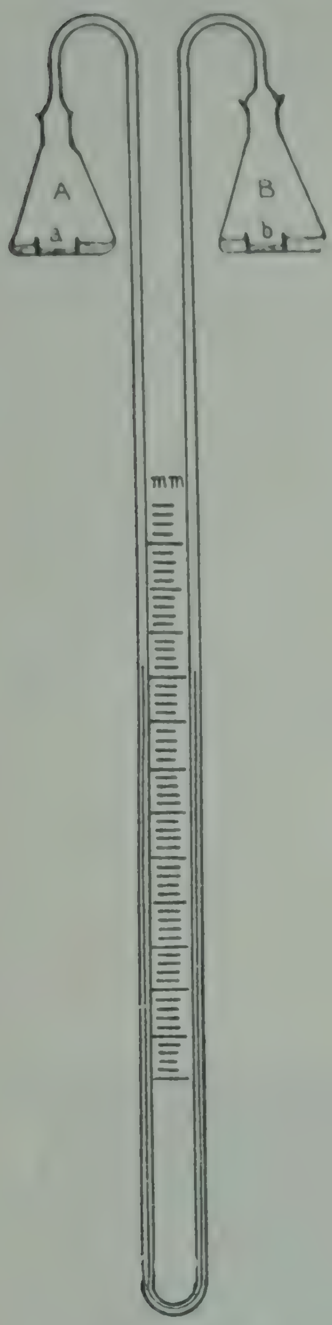


Fig. 69.

If a parallel experiment is made with omission of the potash from the receptacles the change of volume is due to the difference between the oxygen used and the carbon dioxide formed, and from this, if the oxygen used is known, the respiratory quotient can be calculated.

If a thin muscle be placed in oxygen in a microrespirometer, its resting oxygen usage is found to be quite small, and its R.Q. is 0·85 to 0·90.*

When muscle is stimulated, its oxygen usage and all its chemical changes are accelerated. If poisoned with sodium azide, NaN_3,† or cyanide, the oxygen utilisation is suppressed.

If the muscle is stimulated in nitrogen, or after cyanide, lactic acid very quickly accumulates, and the "fatigue maximum" is soon reached. If the muscle is a thin one, *e.g.* the frog's sartorius, and it be suspended in oxygen-free Ringer's solution, the lactic acid diffuses away and fatigue does not set in so soon.

When oxygen is readmitted to a muscle which has been brought to fatigue in nitrogen, the oxygen usage is raised above the former resting value, while the "oxygen debt" contracted during the anaerobic period of stimulation is being wiped off ; this corresponds with removal of lactic acid and recovery of excitability of the muscle. This increased oxygen usage ("excess oxygen") following activity is at first high ; it subsides, at first quickly, then more slowly, to the original basal level, which may only be reached after many hours. As soon as the oxygen usage has returned to the resting level, but not before, the lactic acid is also found to have the low value which characterises the resting tissue. The "excess CO_2," is found to be equal to the excess oxygen, so that the respiratory quotient of the recovery process is unity ; that is to say, it suggests that the oxygen which is utilised in recovery burns up either lactic acid or carbohydrate.

But when we compare the excess oxygen usage in recovery with the *amount* of lactic acid which has disappeared it at once becomes clear that the whole of the lactic acid cannot have been removed by oxidation, for, according to the equation

$$\begin{array}{llll} C_3H_6O_3 + & 3\,O_2 = & 3\,CO_2 + & 3\,H_2O \\ (90\text{ g.}) & (67\text{ l.}) & (67\text{ l.}) & \end{array}$$

each mg. of lactic acid removed should require 744 c.mm. of oxygen for its oxidation, whereas under usual conditions a usage of only 149 c.mm. of oxygen per mg. of lactic acid lost is an average figure, that is, about one-fifth of the amount which would be needed for complete oxidation of the lactic acid that has gone.

The fate of the lactic acid which is *not* oxidised becomes clear when the glycogen content of the muscle is followed during recovery, for there is an increase in the glycogen content of the muscle which is equivalent to the amount of lactic acid removed, but not oxidised. We can thus write the following equations for the separated anaerobic and recovery stages concerned in the formation and subsequent removal of 1 g. of lactic acid, assuming recovery efficiency of one-fifth.

---

* Fenn. *J. Cell. and Comp. Physiol.*, 1932, **2**, 233.

† Stannard. *Amer. J. Physiol.*, 1939, **126**, 196.

(1) $\underset{\text{Glycogen.}}{C_6H_{10}O_5} + H_2O = \underset{\text{Lactic acid.}}{2\,C_3H_6O_3}$

$162\text{ g.} \longrightarrow 180\text{ g.}$

$\therefore\ 0{\cdot}9\text{ g.} \longrightarrow 1\text{ g.}$ (O_2 usage missed in anaerobic period = 149 c.c.)

Anaerobic period

(2) 1 g. lactic acid → 0·2 g. oxidised

$\underset{(0{\cdot}2\text{ g.})}{C_3H_6O_3} + \underset{(149\text{ c.c.})}{3\,O_2} = \underset{(149\text{ c.c.})}{3\,CO_2} + 3\,H_2O$

→ 0·8 g. converted to glycogen (0·72 g.)

$\underset{(0{\cdot}8\text{ g.})}{2\,C_3H_6O_3} = \underset{(0{\cdot}72\text{ g.})}{C_6H_{10}O_5} + H_2O$

Recovery period

The lactic acid cycle now becomes:—

$$\begin{tikzcd}
\text{glycogen} \arrow[d, bend left=0] \\
\text{Lactic acid} \arrow[u, bend right, "\frac{4}{5}"'] \arrow[d] \\
\frac{1}{5}\text{ oxidised to } CO_2 + H_2O
\end{tikzcd}$$

In an atmosphere of oxygen the effect of stimulating the muscle is to raise its oxygen usage to ten or fifteen times the resting value ; if the muscle were thin enough to allow of access of oxygen to all its parts, lactic acid would not accumulate, as other reactions occur and lactic acid is not formed.

OTHER CHEMICAL CHANGES RESULTING FROM CONTRACTION

Associated with the contraction of muscle are a remarkable series of interlocked reactions. The changes most studied are the anaerobic ones, and many of the reactions concerned have been reproduced in cell-free extracts or enzyme solutions. We shall first deal mainly with the anaerobic changes, and later enquire in what way the changes are different when aerobic conditions prevail. Most of the changes too, are recovery changes, *i.e.*, they occur *after* the actual contraction of the muscle, and serve to prepare it for the next contraction. Several of the reactions are coupled ones, *i.e.* two separate reactions run simultaneously, and the occurrence of one of them depends on the possibility for the concurrence of the other. This is often because one of them is associated with a gain, and the other with a loss, of free energy ; the former reaction cannot occur unless the other is proceeding simultaneously and thus able to provide the necessary energy required for the first. Of the substances involved some are individual compounds, and some are teams, or systems.

Among the important changes on contraction are changes in creatine phosphoric acid. After contraction the unstable creatine phosphoric acid rapidly breaks down into creatine and phosphoric acid. Almost immediately after the contraction, even in absence of oxygen, part of the creatine phosphoric acid is resynthesised, and in presence of oxygen it is all rapidly reformed.[*]

Contraction without Lactic Acid Formation. The correlation between muscular contraction and lactic acid formation led, very naturally, to the view that the initial reaction was the conversion of glycogen into lactic acid, with the corollary that the presence of the lactic acid in the muscle in some way caused

[*] EGGLETON and EGGLETON. *J. Physiol.*, 1928, **65**, 15.

it to contract, or, if in sufficient amount, led to rigor. It is now evident that the lactic acid formation follows the contraction, and is itself a sequel to earlier changes. This, while adding further detail, in no way invalidates what had been discovered concerning the lactic acid cycle.

This change of viewpoint arose in consequence of the discovery by Lundsgaard in 1930 * that it is possible to get contraction without lactic acid formation, in muscles poisoned by monoiodoacetic acid. If a frog's muscle be soaked in Ringer's solution containing 1 : 20000 of sodium iodoacetate it can still respond by contraction to stimulation like a normal muscle. When it has given, say, 150 contractions, however, it is fatigued. Whereas in unpoisoned muscle this would lead to the formation of about 0·05 per cent. lactic acid, in the poisoned muscle there is no increase of lactic acid at all. The poisoned muscle, once fatigued, fails to recover when allowed to rest in oxygen, as a normal muscle would do, and soon passes into rigor ; and then, instead of the usual high concentration of lactic acid, there is again found to be no increase over the resting value. The glycogen, however, has disappeared as in unpoisoned muscle.

From these facts it is evident that lactic acid formation is not an essential concomitant of contraction ; yet when lactic acid formation is inhibited the muscle is only capable of a limited number of contractions.

When the creatine phosphate of the muscle is studied after poisoning, it is found that it is broken down after contraction much in the normal manner, but, unlike the unpoisoned muscle, there is no resynthesis of creatine phosphate when the muscle is allowed to rest in absence of oxygen, though there is some restoration in oxygen : moreover, the phosphate set free is converted to a great extent into hexose monophosphate and diphosphate.

A.T.P. and Muscular Contraction. It now seems certain that adenosine triphosphate plays a fundamental *rôle* in all biological processes involving energy liberation, of which muscle contraction is one example. Probably the first event in the onset of contraction is a physical one, as a result of which A.T.P. is brought into contact with an enzyme by which it is broken down to A.D.P. and phosphate,

$$\underset{\text{(A.T.P.)}}{\textcircled{\scriptsize Ad}\text{-HPO}_3 \sim \text{ph} \sim \text{ph}} \longrightarrow \underset{\text{(A.D.P.)}}{\textcircled{\scriptsize Ad}\text{-HPO}_3 \sim \text{ph}} + \text{H}_3\text{PO}_4$$

thereby resolving an energy-rich bond, and liberating about 10 cals. of free energy per mole. The first striking fact is that the enzyme (adenosine-triphosphatase) which does this is normally very closely bound to the muscle protein myosin, *i.e.* the contractile machinery itself. The primary importance of A.T.P. in contraction is further shown by the fact that when applied in minute amounts directly to isolated single fibres of frog muscle, even if curarised, tetanic contractions with action potentials result,† and contraction of mammalian muscle also results on close arterial injection of A.T.P.‡

The second important fact is that the free energy so made available can probably be transformed with little loss into the mechanical strain of the contractile process ; this is almost certainly effected by intramolecular changes, tending to shortening of the long myosin molecules, akin to those which stretched keratin undergoes on contraction (p. 43). Perhaps also the folding of the long molecular chain may be made more rigid by the fugitive appearance of cross-linkages, *e.g.* of the thio-ether type, which are resolved again on relaxation.

Very soon afterwards the A.T.P. is resynthesised again, and the machinery is thus set for a fresh contraction. This resynthesis of course restores the energy-rich bond, and it is done by the transference of the energy-rich bond from creatine phosphate :—

---

* LUNDSGAARD. *Bioch. Ztschr.*, 1930, **217**, 162.

† BUCHTHAL *et al.* *Acta. Physiol. Skand.*, 1944, **8**, 271.

‡ BUCHTHAL and KAHLSON. *Ibid.*, **8**, 317, 325.

$$\mathrm{Cr} \sim \mathrm{ph} + \textcircled{\scriptsize Ad} - \mathrm{HPO_3} \sim \mathrm{ph} \longrightarrow \textcircled{\scriptsize Ad} - \mathrm{HPO_3} \sim \mathrm{ph} \sim \mathrm{ph} + \mathrm{Cr.}$$

The next stage in the normal recovery process must now be the resynthesis of creatine phosphate, with creation of a new energy-rich phosphate bond. This is brought about by pyruvic acid formation, and other later changes, by which energy is set aside for this synthesis. In the formation of 2 molecules of pyruvic acid from glycogen four energy-rich bonds are created in the form of a synthesis of A.T.P. from A.D.P. (*v.* reactions 11 and 14 in table, p. 144).

One of these new molecules of A.T.P. is used up to regenerate creatine phosphate by the reversal of the previous reaction :—

$$\mathrm{Cr} + \textcircled{\scriptsize Ad} - \mathrm{HPO_3} \sim \mathrm{ph} \sim \mathrm{ph} \longrightarrow \mathrm{Cr} \sim \mathrm{ph} + \textcircled{\scriptsize Ad} - \mathrm{PHO_3} \sim \mathrm{ph.}$$

The high free energy content of A.T.P. is therefore in the long run derived (under anaerobic conditions) from energy interchanges connected with the formation of pyruvic acid ; the immediate resynthesis of A.T.P. is ensured by transfer of the phosphate group from creatine phosphate ; it is only later on, during anaerobic recovery, that A.T.P. is regenerated by the slower processes connected with pyruvic acid formation, and then, when there is a piling-up of A.T.P., the creatine phosphate is synthesised again by reversal of the reaction by which it was formed.

The energy-rich phosphate bond can be likened to money. When in the form of A.T.P. it resembles actual cash in hand ready for instant use. This cash can be quickly replaced by drawing at once on a current account at a bank (creatine phosphate). The amounts taken from the bank can be restored by earning money by expending energy in doing work (pyruvic acid formation): but this may be slow and laborious.

It would seem that monoiodoacetic acid checks the formation of pyruvic or of lactic acid by interfering with the breakdown of the triose phosphate. Fluorides poison the conversion of phosphoglyceric acid to phosphopyruvic acid.

HYDROGEN-ION CHANGES DURING CONTRACTION. Muscle is provided with several types of buffer, by which its reaction, normally pH 6·9, is prevented from changing rapidly during contraction.

First, we have the products formed when creatine phosphate breaks up : the base (creatine) then liberated at first causes a slight rise in pH, which reaches a maximum just after relaxation.* Under conditions of free oxygen supply little if any lactic acid is formed, and $\mathrm{CO_2}$ produced is mainly carried away, the excess being dealt with by the muscle buffers, such as salts of proteins. When oxygen supply is less abundant, lactic acid is formed, mainly after relaxation has occurred, and then, in spite of these buffers the hydrogen ion concentration of the muscle does increase after prolonged activity, though not greatly. Then, in prolonged activity we have traces of ammonia liberated by the deamination of adenylic acid (derived from A.T.P.).

FORMATION OF HISTAMINE. It has been claimed that during contraction, and also under the influence of curare, histamine is liberated from a pre-existing store in the muscle and may be detected in the venous blood from muscles.† The amount set free by curare is large and generally admitted, but the liberation of histamine by contraction is small, and some deny that any is formed.

LACTIC ACID CYCLE IN THE INTACT BODY

When we turn to experiments on the whole animal (or on man) we obtain some confirmation, but also some modification, of the views derived from experiments on isolated frogs' muscles.

The total energy of the body is derived ultimately from the oxidation of foodstuffs. The products of oxidation are excreted partly with the urine, partly with the expired air, so that examination of these, combined with a measurement of the oxygen taken into the body, serves as a measure of the total oxidations. Many experiments on the effect of exercise have shown that the excretion of nitrogenous substances is not immediately affected,* while on the other hand any change in the amount of muscular activity is reflected at once in a corresponding alteration in the respiratory exchanges. When a man runs he becomes "out of breath," and this is associated with an increased intake of oxygen and excretion of CO_2. We must conclude, therefore, that the main source of muscular energy is the oxidation of carbon, either in the form of carbohydrate or of fat.

---

* D. K. HILL. *J. Physiol.*, 1940, **98**, 467.

† ANREP *et al.* *J. Physiol.*, **96**, 1935, **85**, 409 ; 1939, 240.

A man at rest absorbs from the respired air about 300 c.c. oxygen per minute, and gives out about 250 c.c. CO_2. The quotient of the volumes of CO_2 given out by oxygen absorbed is called the "respiratory quotient," and, when measured over sufficiently long periods, its magnitude depends on the nature of the substances undergoing oxidation. If these were entirely carbohydrate (or lactic acid) the respiratory quotient CO_2/O_2 would be equal to 1 :

$$C_6H_{12}O_6 + 6O_2 = 6CO_2 + 6H_2O$$

If fats alone were being utilised, the respiratory quotient would be 0·71; with proteins alone, it would be 0·81. In man, on a mixed diet, the respiratory quotient at rest is somewhere about 0·85.

Even the mildest form of exercise increases the intake of oxygen and output of CO_2. Work which raises the oxygen consumption to less than threefold the resting value may be called light or moderate ; hard or heavy work raises it to from 3 to 8 times, and maximal work may for a short period bring it to 10 times the resting level, or even higher. On ceasing exercise the increased pulmonary ventilation rapidly falls, and after a period varying from ten minutes to half an hour the respiration and the oxygen intake have returned nearly to their previous amounts.

When the body is at rest, the circulating blood contains about 15 mg. lactic acid per 100 c.c. This originates from glucose in the circulating blood by *glycolysis* :—

$$\underset{\text{Glucose}}{C_6H_{12}O_6} = \underset{\text{Lactic acid}}{2C_3H_6O_3}$$

Glycolysis is always occurring in normal blood, but is accelerated if, as by deep breathing, the blood carbon dioxide is reduced and the blood thereby rendered more alkaline. It is also accelerated as the blood traverses the lungs, the tissue of which exerts a glycolytic action.† Lactic acid is also being removed constantly from the blood as it is formed ; a considerable amount is removed and utilised by the heart,‡ a little may re-enter muscle and there be converted to glycogen, but the largest part of it is taken up by the liver and converted again into glycogen, probably by a process similar to, but quicker than, that which takes place in the muscle. The glycogen formed in the liver can then be converted into blood glucose, from which the muscles form glycogen more readily than from lactic acid. Hence, when at rest, the blood lactate, which represents a balance between formation and removal, remains constant.

In the living body the richness of the capillary supply, by offering oxygen freely, gives favourable conditions for complete oxidation of the pyruvic acid, as shown above. When exercise is taken, unless this be brisk, there is no increase in the lactate of the blood,§ because no lactate is formed in the muscles. When the exercise is violent, however, the oxygen supply is inadequate, and lactic acid is set free. This easily happens under experimental conditions in mammals; within half a minute after a short tetanic contraction, the venous blood coming away from the muscle contains lactic acid at a concentration almost equal to that in the muscle, and very soon the arterial lactate also reaches a high value ||

---

* Wilson. *J. Physiol.*, 1932, **75**, 67.
† Evans, Hsu and Kosaka. *J. Physiol.*, 1934, **82**, 41.
‡ Evans, Grande and Hsu. *Q. J. Exp. Phys.*, 1935, **24**, 347.
§ Owles. *J. Physiol.*, 1930, **69**, 214.
|| Eggleton and Evans. *J. Physiol.*, 1930, **70**, 269.

(see Fig. 70). A rise of blood lactate also suggests that under normal conditions in man some lactic acid is produced in violent contractions and passes out from the muscle into the blood stream. During the subsequent recovery period the blood lactate rapidly falls again owing to removal by liver, heart and muscles, and for upwards of an hour by excretion in the urine.

It will be seen that when the body as a whole is considered, the net result of exercise and recovery therefrom is the same as in the isolated frog muscle, namely, glycogen is burned, and any lactic acid formed is subsequently removed; but whereas in isolated frog muscle the removal of lactic acid is very slow, and is accomplished by combustion and reconversion (of four-fifths) into glycogen, removal of lactic acid from the intact body after violent exercise is relatively rapid, and is effected by reconversion into glycogen in the liver.

When exercise can be maintained indefinitely, the processes of recovery, according to this theory, are keeping pace with the processes of contraction, *i.e.* any lactic acid that is formed during contraction is being steadily removed and replaced as glycogen, so that there is no great accumulation of this substance either in the muscles or in the blood. Indeed, in such exercise as walking at 4·0 to 4·5 miles per hour there is no increase of blood lactate at all. The glycogen is replenished from the blood sugar, as it is used up.

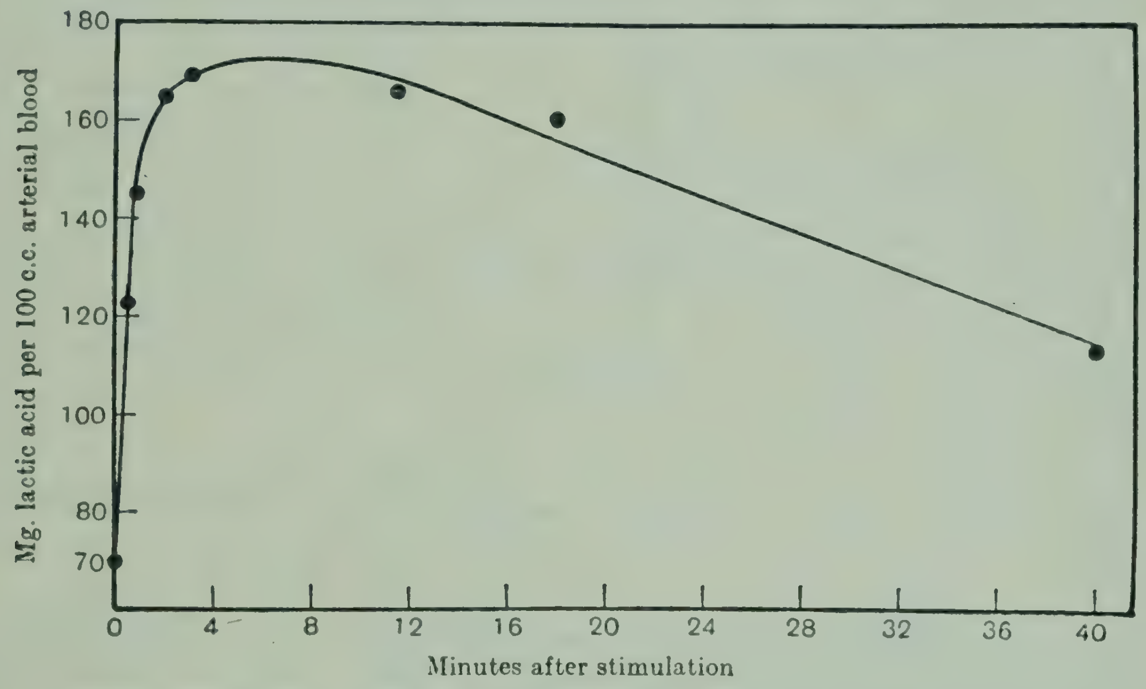


Fig. 70. Lactic acid of arterial blood of anæsthetised dog following powerful muscular contractions resulting from electrical stimulation of legs and trunk. (M. Grace Eggleton and Lovatt Evans, *J. Physiol.*)

When exercise commences the oxygen consumption is accelerated. It may be supposed that the more rapidly pyruvic acid is formed in the muscle, the faster becomes the oxidative removal. Ultimately (assuming the exercise to be moderate) the speed of the recovery process becomes equal to that of the breakdown process.* The oxygen intake per minute rises from the resting level, finally to reach a uniform rate which can be kept up for a considerable length of time (Fig. 71, A and B). This uniform rate is called the "*steady state,*" and signifies the attainment of a steady rate of oxygen usage, together with the upkeep of a constant, but probably quite small, content of pyruvic acid in the muscle by balance of the rates of formation and removal. The more severe the exercise, the higher will be the oxygen usage and pyruvic acid content when the steady state has been reached (Fig. 71). Pyruvic acid, however, does not escape into the blood stream at any time, except in minute amounts, because if the oxygen supply is adequate it is oxidised to CO_2 and water, and if inadequate it is reduced to lactic acid. With normal men taking such exercise as leads to an oxygen usage up to 1·8 l/min. (*i.e.* up to 6 or 7 times the resting intake), as by walking at speeds up to $4\frac{1}{2}$ miles per hour, oxygen usage is adequate either to prevent lactate formation, or to ensure removal of lactate from the muscle without any escape into the blood stream. Whether there is some increase in lactic acid in the muscle, however, is uncertain.*

---

* Hill, Long and Lupton. *Proc. Roy. Soc.*, 1924, **96B**, 438.

As the severity of the exercise is increased the supply of oxygen begins to lag behind the requirement, and some lactic acid begins to be formed in the muscle, some of it overflowing into the blood stream. But so long as the oxygen (blood) supply to the muscles is not grossly inadequate, the lactic acid remains low, so that the total amount present in muscle or blood, and the corresponding oxygen debt, are not excessive. But if the exercise becomes really severe the oxygen supply becomes seriously inadequate and cannot be further increased because the heart cannot pump blood any faster; then the steady state is no longer possible, and lactic acid accumulates more and more, rises to a high level in the blood, and finally brings the performance to an end. In such cases the oxygen debt is a very heavy one, and the process of recovery is prolonged (Fig. 71, curve C, and Fig. 72, curve C).

The maximum severity of this exercise depends on the training of the individual, and physiologically on the extent to which his circulatory system and lungs can provide the working muscles with oxygen in proportion to their needs. The maximum intake of oxygen which, even in an ordinary athlete, can be carried by the blood and pumped round by the heart to the muscles is about 4,000 c.c. per minute. Much severer forms of exercise are, however, possible for brief spells than could be met by such an oxygen supply. In the ten or eleven seconds occupied in a 100 yards sprint, there is not time enough for the oxygen intake to be raised much above the resting level. Immediately after completing the course the man pants violently and in the next few minutes will take in an excess of, say, 4 litres of oxygen. During these eleven seconds the man's muscles have been contracting practically anaerobically, with considerable production of lactic acid. The oxidative removal of this lactic acid, whether in muscle or elsewhere, has been postponed to the end of the race. The man has thus "gone into debt" for oxygen, and the extent to which he can run into debt depends on the amount of lactic acid which can accumulate in the working muscles without stopping their activity. After any severe form of exercise this oxygen debt is apparent (Figs, 71, C; 72). The lactic acid which has accumulated in the muscle during the severe bout of exercise continues to escape when the exercise comes to an end, so that for the first few

Fig. 73.

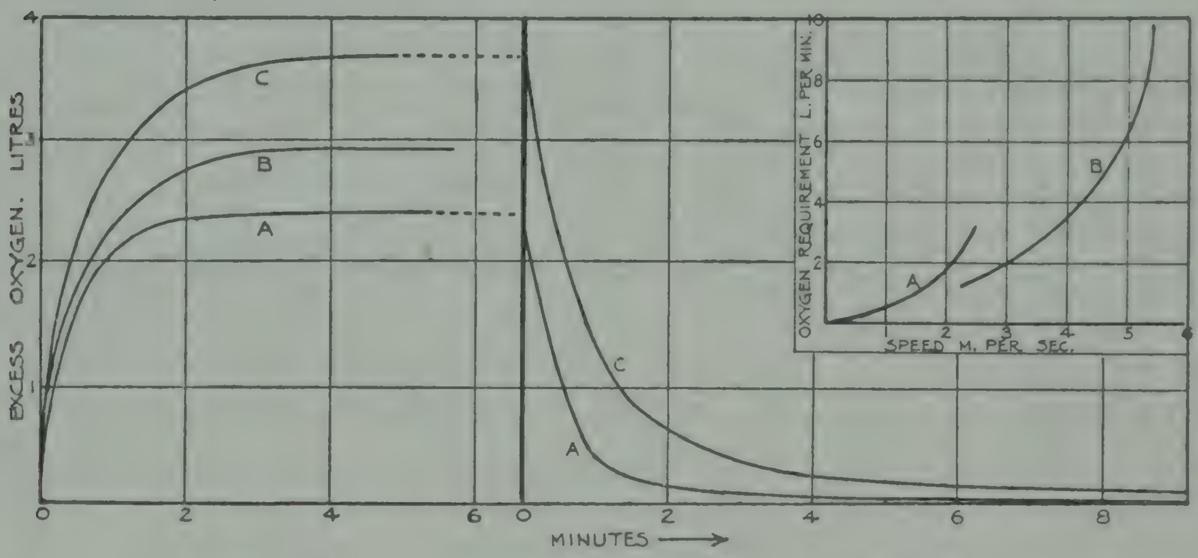


Fig. 71. Fig. 72.

Fig. 71. Excess oxygen intake at commencement of work of different degrees of severity. In A and B the steady state is reached. In C the limit of supply is reached. (After Hill and Lupton.)

Fig. 72. Recovery excess intake of oxygen after work. A and C represent recovery from work comparable with that in the curves A and C of Fig. 71.

Fig. 73. Relation between speed of locomotion and oxygen requirement. Curve A, walking; curve B, running. (After Hill, Long and Lupton.)

---

* Flock, Ingle and Bollman. *J. Biol. Chem.*, 1939, **129**, 99.

minutes its percentage amount in the blood first rises and then falls steadily as the lactic acid is taken up by the liver and muscles and converted into glycogen again.

In spite of the apparent simplicity of the case as presented above, there are certain facts which point to its not being in reality quite so simple. Thus injection of lactate into the blood stream should cause increased oxygen usage, but does not. The calculations based on blood lactate in exercise are thus probably misleading except in so far as they provide a guide to muscle lactate concentrations as an index of extensive but incomplete breakdown of glycogen, and other components of muscle. In all probability very little blood lactate returns to the muscle, most of it being taken up by heart and liver. Hence the failure of attempts to show a close correlation between oxygen usage and fall of blood lactate after exercise.

No doubt the increased oxygen usage known as oxygen debt represents an active oxidation of some component, by which is provided the energy needed to restore the potential energy of the muscle to its resting value by making good the energy-rich bonds resolved by contraction. But we do not know whether lactic acid is the substance, or the only one, so oxidised.

Effect on Respiratory Quotient. Though moderate exercise does not alter the respiratory quotient, very considerable disturbances are brought about by severe exercise. These are due to driving out of CO_2 during the exercise, followed by compensatory retention of CO_2 afterwards.

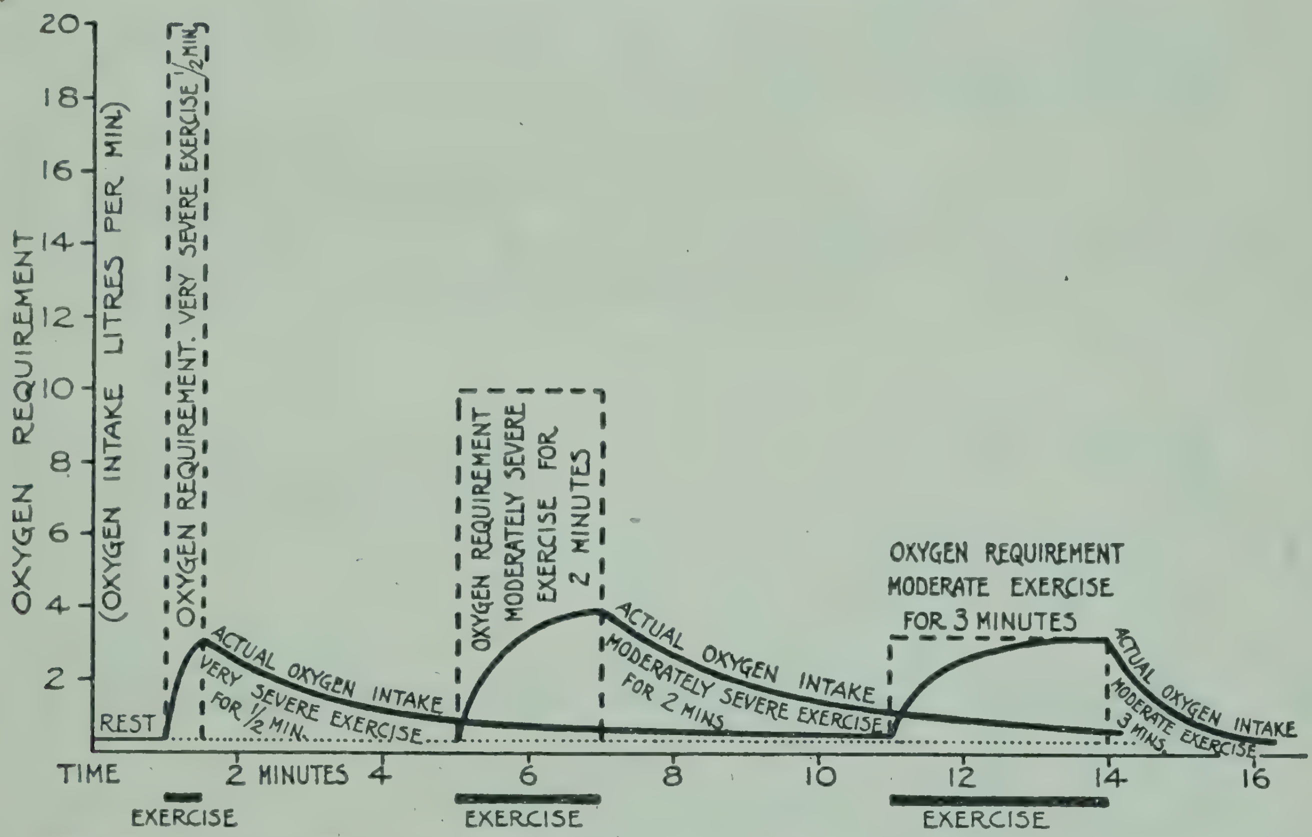


FIG. 74. Diagram to illustrate the Meaning of "Oxygen Requirement," and the Distinction between "Oxygen Requirement" and "Oxygen Intake."

The area of an oxygen requirement rectangle is, by definition, the same as that of the complete oxygen intake curve, each being measured from the level of the resting oxygen intake (shown dotted). Three periods of exercise of different severity.

During exercise, there is considerable stimulation of the respiratory movements. Such a stimulation is necessary for the increased intake of oxygen, but it also simultaneously, and to a greater degree, quickens the elimination of CO_2. Moreover, the ultimate escape of lactic acid into the blood, when the exercise is violent, will drive out CO_2 and will further increase its elimination by the lungs. Thus an amount of carbon dioxide would be blown off from the lungs in excess of that produced by the metabolism of the body. For example, if the true metabolism during exercise were: oxygen 3 litres per minute, CO_2 2·4 litres per minute, the true respiratory quotient would be 0·8. If, however, 3 litres extra of carbon dioxide were displaced, and this was eliminated in ten minutes, the apparent respiratory quotient would be raised to 0·9. This rise of CO_2 output and of respiratory quotient would be the initial event in any steady form of exercise, so that the increased breathing would be more marked while this extra CO_2 was being eliminated. After its elimination the organism enters into a steady state: the excessive breathlessness disappears, the individual having attained "second wind." With severe exercise of short duration, when a considerable oxygen debt is incurred, the escape of lactic acid into the blood causes a great dyspnœa immediately after the exercise, with a large output of CO_2. The apparent respiratory quotient under these circumstances may rise to 1·5 or even to 2. During the process of recovery the lactic acid is gradually removed, the base by which it was neutralised is liberated, and the excess CO_2 which has been given off from the body is once more retained. We therefore find that the high respiratory quotient immediately after heavy exercise is followed later on by a long lasting fall of the respiratory quotient below the resting level.

Substances Consumed by Working Muscle. For brief exercise of moderate severity it has been argued that if we make due allowance for these fluctuations in the respiratory quotient, due to the initial driving off of the CO_2 and its subsequent retention as the lactic acid disappears, we arrive at the corrected respiratory quotient of the exchanges, directly due to muscular exercise of short duration. In making this calculation it is customary to deduct the oxygen usage and CO_2 production during rest and to call the remainder the "excess metabolism" of exercise plus recovery. One of the difficulties in this type of experiment is that of deciding how long a time is necessary for return of the respiratory exchange to the normal resting level. It has been claimed that this excess metabolism has a R.Q. of about 1·0. We are therefore perhaps entitled to conclude that in the whole animal, as in the isolated frog's muscle, the energy of muscular contraction, when this is of moderate severity, is normally derived ultimately from the oxidation of a carbohydrate, to CO_2 and water. In support of this, too, is the fact that during exercise the blood sugar is often lowered, presumably by acceleration of uptake by the muscles.

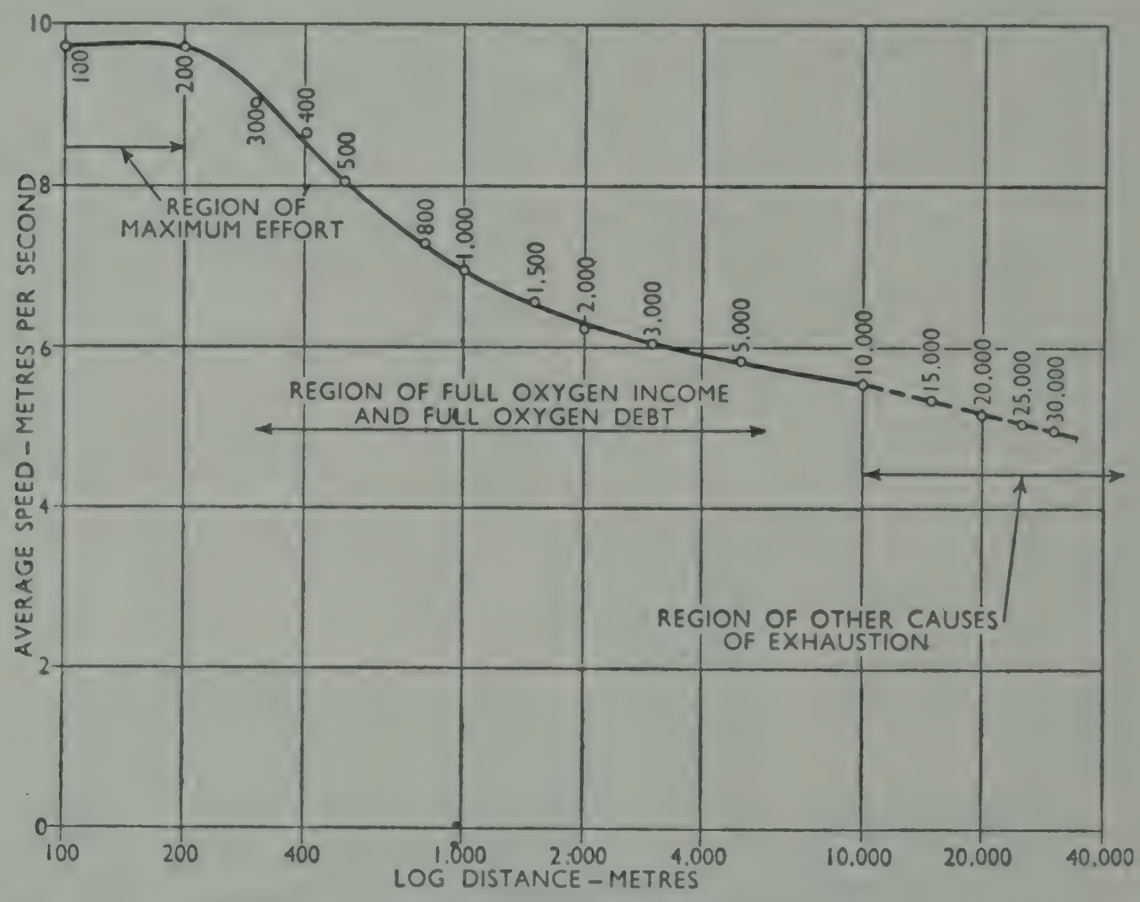


Fig. 75. Athletic records. Average speed plotted against log distance. (A. V. Hill.)

That carbohydrate, though the choice, is not the sole possible source of muscular energy, however, is indicated by several converging lines of evidence. The carbohydrate stores of the muscles can be brought to a low level by fasting, or by a diet consisting almost entirely of fat, and containing no carbohydrate. After three days of such a diet the subject is disinclined for exercise and easily susceptible to fatigue. Ketones derived from fat metabolism are present in blood and urine; when exercise is taken the amount of these is reduced instead of the blood sugar being lowered. After the exercise the amount of ketone bodies increases still further, suggesting that the liver has furnished them at an accelerated rate.*

It has been shown both for man and dog that prolonged and violent exercise can be taken in which the material consumed amounts to ten times the carbohydrate content of the entire body.† Whether under these circumstances fat or protein is consumed direct or whether these are first converted into carbohydrate is uncertain, but it would seem most probable that it is the latter.*

---

* Neufeld and Ross. *Amer. J. Physiol.*, 1943, **138**, 747.

† Dill, Jones and Edwards. *Le travail humain*, 1934, **2**, 1.

With very intense exercise the R.Q. of the excess metabolism may be very high (up to 1·7), a result as yet unexplained.†

THE "MECHANICAL EFFICIENCY" OF HUMAN EXERTION. It is possible to calculate the amount of mechanical energy liberated (theoretical maximal work) by men carrying out various exercises under normal conditions, and then by determining their oxygen utilisation during the work and recovery, to compute the mechanical efficiency of the performance. This is, under favourable conditions, about 20 per cent.‡ Many approximations have to be made and sources of error allowed for, but it is surprising with what constancy an experienced athlete can perform. In running, which has been chiefly studied, the allowances to be made include those for starting up and stopping and for onset of fatigue. In running, little actual mechanical work is done, the energy set free being used for imparting velocity to the body, a good deal of it being lost in overcoming viscous or other resistance in the muscles and other structures. The speed reached depends on the ratio of propelling force to resistance, chiefly that of his own muscles, as viscosity, and that represented by starting and stopping the movements. The resistance of the air is normally small, and may be allowed for. The superiority of first-class athletes over others of equal muscular strength appears to be due to the smaller "viscosity" of their muscular apparatus, using this term in its widest sense to mean that in them less force is wasted in overcoming the resistances represented by putting their limbs in motion.

In maximal performances, such as those involved in athletic records, the man is exhausted when his oxygen debt reaches, say, 15 l., while his maximal oxygen intake is usually about 4 l./min. Lash, who created a world's record for a 2-mile run, used 5·35 l./min. when running at 21·6 km./hr.: in runs at 18·7 km./hr. for 5 min. he reached 5·1 l./min. and at the finish had only about 48 mg./100 c.c. of blood lactate.§ Hence a man can do work at a greater rate the shorter the total time for exhaustion, so that there is a close relation in athletic records between distance covered by a runner at exhaustion and the speed at which he covers it, as shown by Fig. 75.

During really maximal exertion much of the work is performed anaerobically, and lactic acid is produced at about 3 to 4 g./sec., and as a maximal tolerable amount is about 90 grammes, it is calculated that the severest exertions can only last for about 30 sec. This agrees with the practical experience of athletes.

For Reference

CHRISTENSEN, KROGH and LINDHARD (1936). Studies of Severe Muscular Exercise. *Skand. Arch. Physiol.*, **74**, 1.

CORI (1945). Enzymatic Reactions in Carbohydrate Metabolism. *Harvey Lects.*, **41**, 253.

GEMMILL (1942). The Fuel for Muscular Exercise. *Physiol. Rev.*, **22**, 32.

HILL (1932). The Revolution in Muscle Physiology. *Physiol. Rev.*, **12**, 56.

MILLIKAN (1939). Muscle Hæmoglobin. *Physiol. Rev.*, **19**, 503.

NEEDHAM (1932). The Biochemistry of Muscle. (Methuen, London.)

SACKS (1941). Changing Concepts of the Chemistry of Muscular Contraction. *Physiol. Rev.*, **21**, 217.

THE HEAT PRODUCTION IN MUSCULAR CONTRACTION

Although chemical investigation has yielded valuable information as to the various stages by which net chemical changes occur, it does not give full information as to the course of events occurring during contraction, relaxation and recovery. All we can do is to examine the summated results of a long series of contractions or to study the action of individual enzymes. A single twitch will give no detectable chemical change, yet energy has been set free, and part of this appears in the form of heat. Methods for detecting small differences of temperature are now so exact that we are able to measure not only the heat production during a single twitch but also the time-course of its evolution.

In large animals the production of heat in muscular contraction can be easily shown by a thermometer. This observation is confirmed when we investigate the contraction of an isolated muscle outside the body. If a frog's

---

* GEMMILL. *Bull. Johns Hopk. Hosp.*, 1940, **66**, 71.

† BEST, FURUSAWA and RIDOUT. *Proc. Roy. Soc.*, 1929, **104B**, 119. SOLANDT and RIDOUT. *Proc. Roy. Soc.*, 1933, **113B**, 327.

‡ DICKINSON. *J. Physiol.*, 1929, **67**, 242.

§ v. SCIENCE, 1937, **85**, 409.

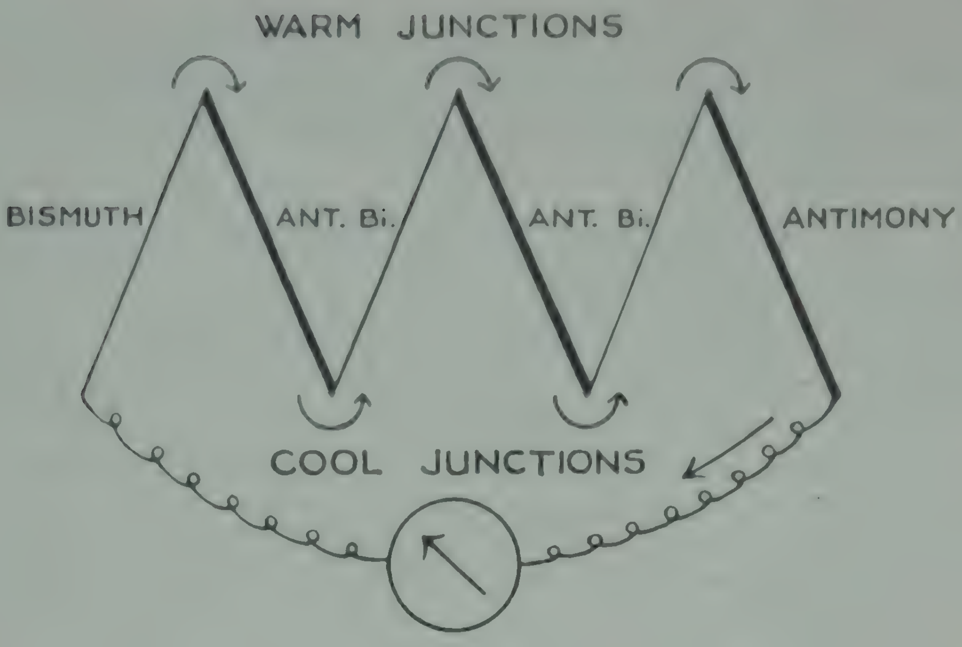


FIG. 76.

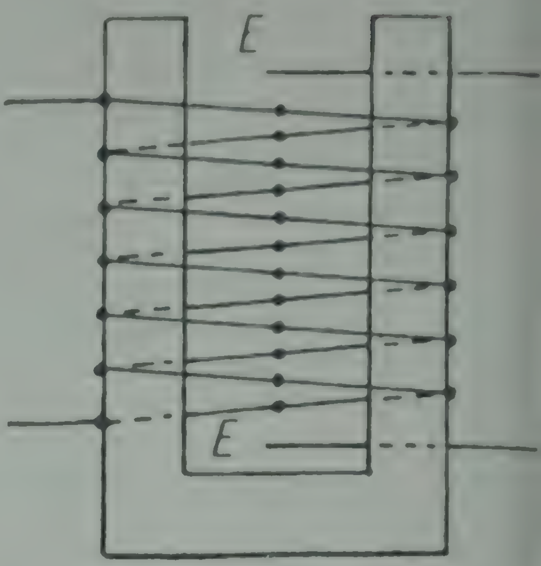


FIG. 77. Plan of Construction of Thermopile. (HARTREE.)

muscle at 20° C. is tetanised, its temperature rises about 0·03° C. per second o stimulation, and for each single twitch from 0·001° to 0·005° C.

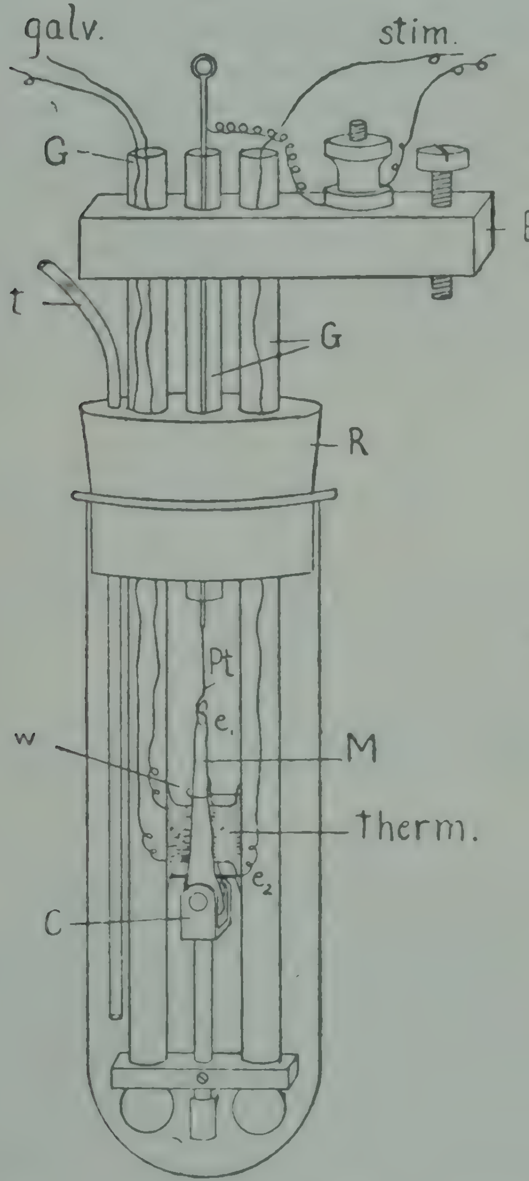


FIG. 78. Arrangement of Thermopile and Muscle for measuring Heat Production during Contraction. E is an ebonite block by which the whole apparatus is secured upright in the water bath; *w* is a platinum wire guide for the muscle to insure good contact with the thermopile; *t* is a fine glass tube through which Ringer's solution or oxygen can be introduced into the chamber; C is an ebonite clamp to hold the muscle; it is secured to the glass rods by means of an ebonite bar at the base and a silver-plated set screw. (W. O. FENN, *J. Physiol.*, 1923, **58**, 175.)

These small changes in temperature can be estimated by thermoelectric junctions. When two or more junctions of a circuit made of two metals are at different temperatures, a current of electricity flows through the circuit, and can be measured by means of a galvanometer. Thus in the circuit (Fig. 76) composed of antimony and bismuth, if the upper junctions be warmed there will be a current flowing in the direction of the arrows, and this current will within limits be proportional to the difference of temperature between the warm and cool junctions. In such a series of junctions, alternately warmed and cooled, the currents are added together; this is the thermopile.

To measure the rise of temperature in a muscle during its contraction, the thermopile is made in the form of a grid. It consists of forty or more turns made with alternate segments of fine iron and constantan wire soldered together with silver solder and wound between the arms of a U-shaped insulating frame (Fig. 77) so that two sets of junctions lie half-way between the arms of the U and the other two sets on the outside edges of these arms. A slip of mica is inserted between the front and back parts of the spiral, and the wire insulated throughout wit a coating of shellac varnish. A pair of muscle (*e.g.* sartorii) is used, and is arranged so as to straddle the grid, each muscle lying close against the middle ("warm") junctions, while the oute junctions are maintained at a constant temperature. For this purpose the arrangement show in Fig. 78 may be used. Through the rubbe bung, R, of a large test tube pass three glass tubes G. Between these the thermopile is fastened The leads from the thermopile to the galvanomete pass through two holes in one of the glass tubes The muscle is stimulated by the electrodes, e and e_2. The lower electrode e_2 is connected wit the stimulating circuit by means of a copper wir leading up through the other glass tube. Th upper electrode e_1 consists of an 8-shaped loo of platinum tied between the two muscles, an soldered to a fine steel rod, which makes electrica connection with the stimulating circuit, an mechanical connection with the lever whic records the contractions of the pair of muscle The whole apparatus is kept at as constant temperature as possible. The leads from the thermopile are taken to a very sensitive galvanometer of low resistance.

The thermopile and galvanometer can be calibrated so that the rate of heat production by the muscle may be deduced directly from the galvanometer excursions. For this calibration, the muscle at the end of the experiment is killed, and the heating effect observed on passage of an alternating current. Knowing the strength of the current and the resistance of the dead muscle, the exact amount of heat required to produce an observed deflection can be calculated.*

Calorimetric Method. For the determination of the total heat production in a muscle or group of muscles over a considerable time we may use microcalorimeters.†

Two small Dewar flasks are used (Fig. 79). In each of these are suspended several frogs' muscles immersed in Ringer's fluid and arranged so that they can be stimulated. Thermojunctions are immersed in the fluid in the two flasks. Any difference of temperature between the two vessels as revealed by the galvanometer is therefore due to liberation of heat in the muscles which are being excited.

Valuable results have been obtained with the above-mentioned methods on the heat production during muscular contraction. By the calorimetric method it was shown that in anaerobic muscular contractions the formation of 1 g. of lactic acid is attended with the evolution of about 370 gramme-calories.‡ The thermopile shows that this heat does not come off at once, but in two phases, there being an evolution of "*initial heat*" immediately following on excitation, and a subsequent evolution of "*anaerobic delayed heat,*" following relaxation.

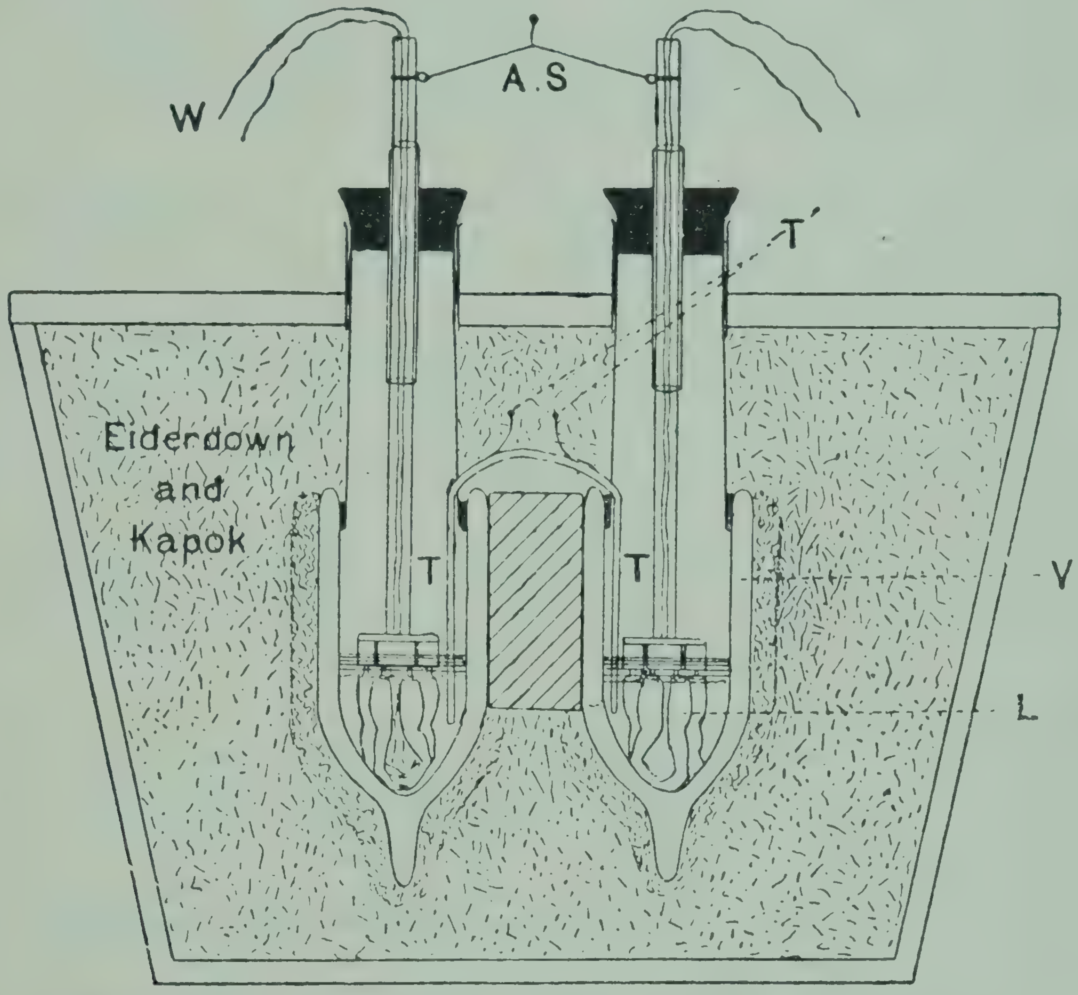


Fig. 79. Arrangement of two Dewar Flasks for studying the Production of Heat in Muscles over a considerable period of time. (R. A. Peters.)

During recovery in oxygen there is, accompanying the disappearance of lactic acid, a further prolonged evolution of heat, "*delayed aerobic heat.*" It has been shown, by cooling frog muscles to 0° so that changes are slowed down, that the time course of the oxidative delayed heat is the same as that of oxygen consumption.§ As this is only a further 370 calories per gramme of lactic acid removed, and since oxidation of 1 gramme of this substance involves the liberation of 3,600 calories, it is clear that either the oxidation of lactic acid to CO_2 and water is incomplete, or else must be coupled with some other reactions which are accompanied by considerable absorption of energy. Meyerhof's explanation was that about one-fifth of the lactic acid undergoes complete oxidation, but that only part of the energy of this oxidation appears as heat, while the rest of it is absorbed in the endothermic reconversion of lactic acid into glycogen.

---

* Hill. *Proc. R. S.*, 1928, **103B**, 117.
† Hill. *J. Physiol.*, 1912, **43**, 261.
‡ Peters. *J. Physiol.*, 1914, **47**, 243.
§ D. K. Hill. *J. Physiol.*, 1940, **98**, 207.

ISOMETRIC CONTRACTIONS. For technical reasons, a great deal of work has been carried out on the heat production in isometric contraction, and it was supposed that the only difference on isotonic contraction was that the tension of the isometric contraction was more or less completely converted into work if the muscle was allowed to shorten. As already explained, however, isometric contraction is somewhat of a myth when considered from the energy standpoint, and moreover, the relation between isometric and isotonic contractions cannot be arbitrarily simplified in the manner stated.

Though much has been learned by study of the isometric contraction, recent improvements in technique have enabled the simpler case of the isotonic contraction to be adequately studied.

The total initial heat formed in isometric contraction and relaxation is a measure of the total energy set free by the muscle in the act of its contraction, and, as Hill showed, the tension set up under isometric conditions bears a constant relationship to the heat production, the ratio Tl/H (where T = tension, l length of muscle, and H the heat produced) remaining constant at an average value of 8, under a variety of conditions. This ratio is a measure of the perfection with which the chemical potential energy is converted into mechanical potential energy, *i.e.* of the efficiency of the muscle.

Increase of initial extension causes a rise of the heat production, as well as of tension production, up to a certain maximum, beyond which there is a decline in both (*v.* Fig. 54).

Isotonic Contractions. It has been shown by Fenn * that a muscle develops more heat on isotonic than on isometric contraction. This may be interpreted as meaning that the isotonic contraction does not merely utilise tension energy liberated at the commencement, as though the muscle were a stretched spring, but that it continues to develop further power as the shortening progresses. In accordance with this it has been shown that the relation between force exerted and rate of shortening is not as simple as a viscous elastic model would suggest, *i.e.* the decrease of force with increase in speed is not linear, as viscous change would demand, but exponential.†

The heat associated with contraction consists of *initial heat* and *recovery heat*, the former occurring in a sudden burst during contraction, and the latter much more slowly during recovery, *i.e. after* relaxation has ended. Recovery heat is much reduced under anaerobic conditions. Heat is also produced in further amount in maintaining a contraction, and is called "*maintenance heat.*" The relations are roughly :—

(1) *Contraction heat*		
(*a*) Shortening heat		1·0
(*b*) Work heat	varies	
(*c*) Maintenance heat	varies	
(2) *Relaxation heat*	varies	
(3) *Recovery heat*		1·0
(*a*) Aerobic		1·0
(*b*) Anaerobic		0·08

The heat production under isotonic conditions is complicated by a *thermo-elastic effect :* resting muscle, like rubber, warms when stretched and cools when it is released again ; contracted muscle does the reverse.

The *initial heat* begins at the same moment as the shortening,‡ starts off at its maximum, and finishes as soon as shortening ends. The amount of *relaxation heat* depends on conditions : if no work is dissipated in relaxation, *e.g.* if the muscle does not have to lower the weight lifted by its shortening, then there is no relaxation heat at all. If the weight does pull on the muscle, then heat is set free in proportion to the load lowered.

---

* FENN. *J. Physiol.*, 1924, **58,** 175, 373.

† FENN and MARSH. *J. Physiol.*, 1935, **85,** 277.

‡ A. V. HILL. *J. Physiol.*, 1948, **107,** 29.P.

The *shortening heat* is proportional to the amount of shortening, but (with maximal stimuli) independent of the load lifted : hence, within a given range, the heavier the load the greater the total energy set free when equal shortenings are considered : but the extra energy appears as work, not as heat. The mechanical efficiency of the contraction may reach 35 per cent.

When a contraction is *sustained*, as in a tetanus, further heat is continuously liberated, and this represents the energy expended in maintaining tension : it is therefore conveniently studied in isometric contraction at different lengths. *Maintenance heat* is increased by increase of initial length of fibre, because this raises the tension that has to be sustained. It is also increased by rise of temperature (2·8 times by 10° C. rise) because tension dissipation is increased by rise of temperature and probably depends on chemical events. The cost of maintenance varies in different muscles and in different species ; it is high in frog as compared with man.

THE MECHANICAL EFFICIENCY OF CONTRACTION. When a muscle contracts it may perform mechanical work, and the energy for this work is derived from the chemical and physical changes associated with contraction and recovery. It is often necessary to consider the *mechanical efficiency* of contractions carried out under certain conditions, *i.e.* to know what fraction of the total energy set free in the contraction is converted into work. Thus if 10 per cent. of the energy set free were converted into work, the remaining 90 per cent. being liberated as heat, the mechanical efficiency would be 0·1 (or 10 per cent.).

The mechanical efficiency will vary, for a given muscle, with conditions, *e.g.* with load lifted, rate of shortening, etc., but will of course never be anything like 100 per cent. In human muscles *in situ* it may be anything up to 20 per cent. and seems to be similar for frog muscle *in vitro*,* when the whole cycle of contraction and recovery is taken into account. For individual muscles speed of contraction largely influences the efficiency, *e.g.* for arm flexion to full extent the optimal time is 0·5 sec., and for a leg flexion about 1 sec. With an isometric contraction, or an isotonic one without any load, the efficiency is, of course, nil.

Mechanical efficiency is measured by finding the work done in the usual way (weight × height) and finding what fraction this is of the energy expended as found by measuring either total heat production, or total oxygen usage. (Knowing that 1 g. cm. = 23·4 micro cal. = $4{\cdot}7 \times 10^{-6}$ c.c. oxygen usage.)

RELATION TO CHEMICAL CHANGE. By analysis of the time relations of the process by the use of thermopiles, and correlation with the chemical changes as discussed above, it would seem that the sequence of events as regards the heat production in muscular contraction and recovery is in outline as follows :

(1) Contraction. The first event is probably the breakdown of A.T.P. into A.D.P. and phosphoric acid, with liberation of heat and energy :—

1 mol. A.T.P. = 1 mol. A.D.P. + H_3PO_4(+ 11,000 cal.).

Then follows the breakdown of creatine phosphate and linked with it the resynthesis of A.T.P. :—

1 mol. creatine phosphate = 1 mol. creatine + H_3PO_4(+ 11,000 cal.).
A.D.P. + H_3PO_4 = A.T.P. (− 11,000 cal.).

(2) Relaxation. Next, three reactions occur simultaneously :—

1 mol. hexose equivalent (glycogen) = 2 mol. lactic acid (+ 32,000 cal.).
2 mol. lactic acid neutralised by proteins, etc. (= + 16,000 cal.).
4 mol. creatine + 4 H_3PO_4 = 4 mol. creatine phosphate (− 44,000 cal.).

and, as will be seen, the net heat exchange is + 4,000 cal.

After relaxation has occurred there is formation of lactic acid and resynthesis of creatine phosphate. There is a further and much slower production of heat.† In actual experiments on muscles, in absence of oxygen this, the delayed anaerobic heat, is much less per g. lactic acid formed than we should calculate. The heat liberated represents the algebraic sum of at least four simultaneous processes :

(*a*) Lactic acid formation—evolution of heat . . . .
(*b*) Lactic acid neutralisation by protein-evolution of heat .
(*c*) Creatine phosphate resynthesis-absorption of heat . .
(*d*) Re-ionisation of protein-absorption of heat . . .

Total = +200 to 300 cal. per g. lactic acid.

---

* HILL. *Proc. Roy. Soc.*, 1939, **127B**, 434 ; **128B**, 263.

† HARTREE and HILL. *J. Physiol.*, 1922, **56**, 367 ; 1923, **58**, 127. BLASCHKO. *J Physiol.*, 1930, **70**, 96.

When the muscle has been poisoned with iodoacetic acid, stage (1) occurs, but not stage (2), and as we have seen, there is no lactic acid formation. Hence it is concluded that the heat available from the formation of lactic acid does not all appear, since some of it is absorbed in the endothermic reactions (*c*) and (*d*) and some perhaps in other ways.

(3) In presence of oxygen the delayed heat is much increased (Recovery heat), though its rate of formation is only $\frac{1}{600}$ as rapid as that of the initial heat.[*] This again is due to a balance of exothermic and endothermic processes :—

(*a*) Complete creatine phosphate resynthesis—(absorption of heat)	total about 370 cal. per gramme L.A. removed.
(*b*) Oxidation of lactic acid (or carbohydrate)—(evolution) .	
(*c*) Re-formation of glycogen (absorption)	

The results of the thermal investigations thus enable us to supplement those obtained by chemical methods but final agreement has not yet been reached.

For Reference

Fenn (1945). Hober's Physical Chemistry of Cells and Tissues, p. 453.

Meyerhof (1937). *Ergebn. d. Physiol.*, **39,** 10.

Moyle-Needham. (1938). Energy-yielding Reactions in Muscle Contraction. *Enzymologia,* **5,** 158.

OTHER PHYSICAL CHANGES ASSOCIATED WITH CONTRACTION

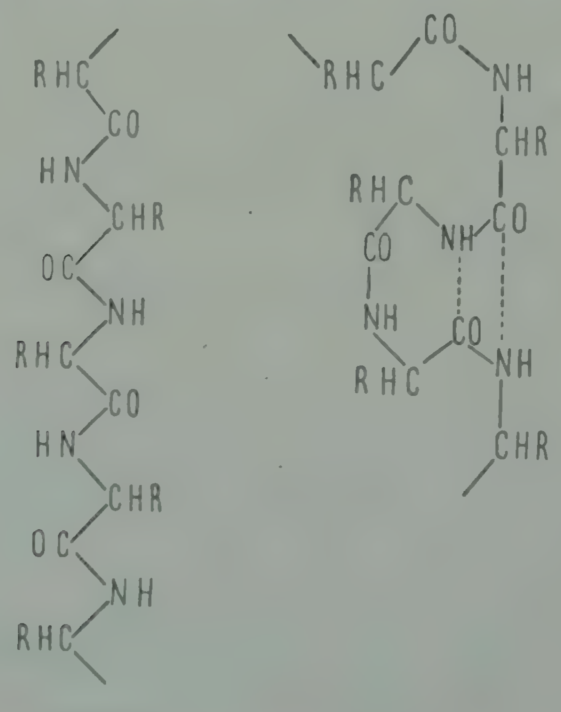
Various other physical changes are associated with contraction ; thus there are changes in double refraction (diminished), volume (decreased), translucency (diminished), and impedance (increased in isometric, decreased in isotonic contraction). Increase of pressure to 300 atm. can cause contraction of muscle.[†] Although these changes have not been very extensively studied, the close correspondence between their time relations and the time relations of the tension curves and heat production indicate that the processes are closely related. Generally speaking, the curves for all these changes show two phases, one during the contraction and the other during relaxation. The former is thought to be related to changes in adenosine triphosphate, and the latter to creatine phosphate and lactic acid changes. In muscles poisoned with iodoacetic acid, or under anaerobic conditions, the time relations of all five changes are also very similar.

THE ULTIMATE CAUSE OF MUSCULAR CONTRACTION

The means by which the chemical energy of the contracting muscle is converted into the mechanical energy of the contractile process is unknown. Accordingly, theories are numerous.

We know now that there can be no question of a conversion of heat energy into work, because the necessary local high temperature could not be reached without killing the muscle.

As a net result, there is no doubt that the source of the energy is oxidative processes. Experiments on the metabolism of the body as a whole, and of mechanically functioning



Resting Contracted

---

[*] Hartree. *J. Physiol.*, 1932, **75,** 273 ; **77,** 104.

[†] Dubuisson. *J. Physiol.*, 1937, **89,** 132.

organs like the heart, show that the performance of work involves an increase of the oxidative processes with a corresponding evolution of energy, of which at least four-fifths will appear as heat while at most one-fifth may be transformed into mechanical work.*

Among theories that have been at various times in vogue are (1) that of attraction between structures of opposite electric charge; (2) that lactic acid liberation caused a physical change akin to coagulation, or that creatine phosphate breakdown was the responsible factor; (3) that increase of surface tension on the surface of, or of osmotic pressure within, some small structural elements, caused the shortening; (4) that it was due to swelling of colloidal particles.

At present opinion is converging upon the view that the ultimate machinery of contraction is the long myosin molecule,† and that in the act of contraction this becomes folded or looped, and so becomes shorter. Since the myosin forms an important fraction of the cytoplasm of the fibre, this leads to shortening of the fibre. The evidence that the myosin molecule is of the long fibrous type is convincing, and it is supposed that in the resting state the molecule is stretched out, as shown above.

The NH– groups have a positive charge and the CO– groups a negative one and so these tend to attract one another; but they are kept at a distance because the molecule is fixed by some subsidiary linkages, *e.g.* by S–S bonds, or by repulsions between other bonds or side-groups, *e.g.* between NH_2 groups of lysine residues. When the muscle is stimulated we may suppose that the enzyme adenosine triphosphatase is set free, or made active; by this A.T.P. is broken down to A.D.P., setting free a phosphate group and about 10,000 cal. of energy per mole of A.T.P. This liberated energy may be in some way responsible for releasing the bonds that keep the myosin molecule in the elongated form, thus allowing full play to the force tending to crumple up the long chain, as shown above. Or the shortening may be brought about by transfer of the phosphate group to the myosin molecule, *e.g.* to the NH_2 group of lysine residues. A liberation of calcium, or of potassium, from a bound form has also been considered a possible factor in releasing the crumpling process, but there is as yet no clear evidence as to what starts the process off.

Whatever the precise mechanism, there is reasonably good evidence, first, that A.T.P. is a prime mover, and second, that the fundamental process is a crumpling of the myosin chains. When the muscle relaxes, these processes are reversed: if any energy is transferred in order to "re-wind the spring" and set the myosin molecule back in the elongated form, it would appear to be transferred quantitatively, without loss as heat, since, as we have seen, relaxation may occur without heat production. But the energy-rich phosphate bonds are, in fact, believed to be transferable without loss. Such a transfer would only be of the nature of a temporary borrowing, to be made good at a slower rate later on. Subsequently, in fact, energy is used, in oxidative reactions, as we have already seen, and the recovery heat is an expression of this energy generation which restores the muscle to its original condition.

For Reference

FENN (1936). Electrolytes in Muscle. *Physiol. Rev.*, **16**, 450.
FENN [ED.] (1941). Muscle (a Symposium). (Lancaster, Pa.: Jaques Cattell Press.)
SZENT-GYÖRGYI (1948). The Nature of Life. (New York: Acad. Press. Inc.)
VERZÁR (1943). Theorie der Muskelkontraktion. (Basel: Benno Schwabe.)

---

* BENEDICT and CATHCART. Muscular Work. (Carnegie Publ., Washington.) 1913.
† ASTBURY. *Proc. R. S.*, 1947, **134B**, 303.

CHAPTER XI

PLAIN MUSCLE AND OTHER CONTRACTILE TISSUES

OUR knowledge of the physiology of unstriated (also called smooth, plain, or involuntary) muscle is derived from experiments on viscera or organs in which it predominates, *e.g.* in mammals, the uterus, arteries, stomach, intestine, ureter, bladder, nictitating membrane and retractor penis, and in special muscles of invertebrates.* Plain muscle often differs from voluntary muscle in containing numerous plexuses of nerve fibres (non-medullated) and ganglion cells, so that it is difficult to be certain whether some results are due to the muscle fibres themselves, or to the nerves and nerve cells which are so intimately connected with them; especially as we have as yet no convenient drug like curare by aid of which we might discriminate between action on muscle and action on nerve.

A further difficulty is that the physiological properties of plain muscle differ with its source, so that it is unsafe to generalise from the results of experiments with tissue from only one source. The tissue occurs in two extreme forms, and there are intermediate types.

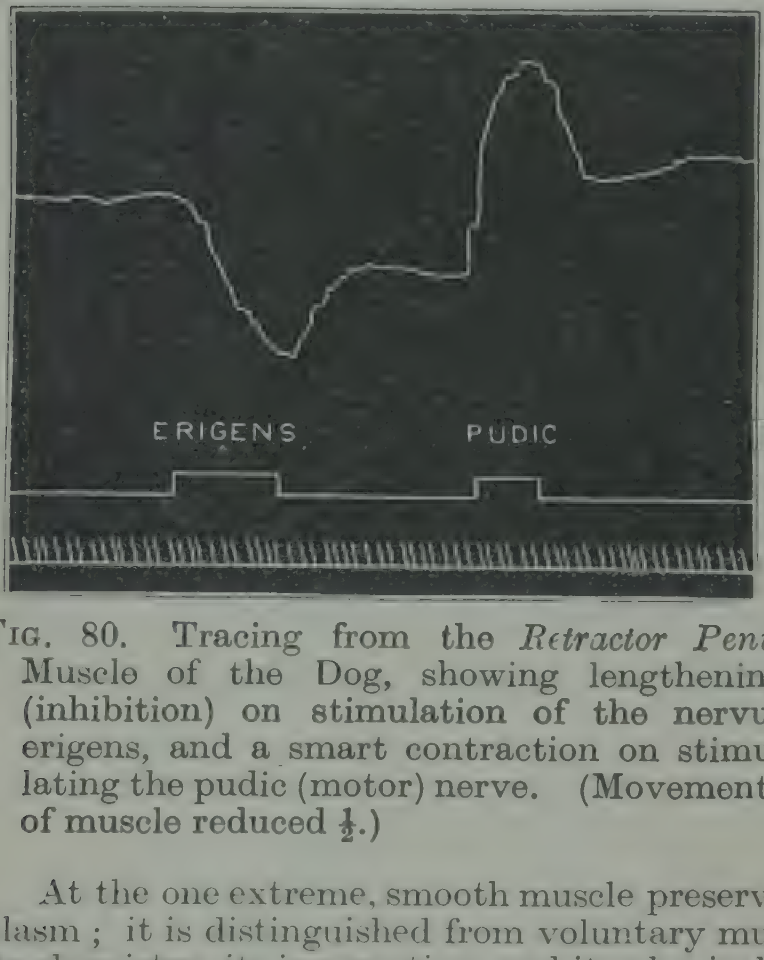


FIG. 80. Tracing from the *Retractor Penis* Muscle of the Dog, showing lengthening (inhibition) on stimulation of the nervus erigens, and a smart contraction on stimulating the pudic (motor) nerve. (Movements of muscle reduced ½.)

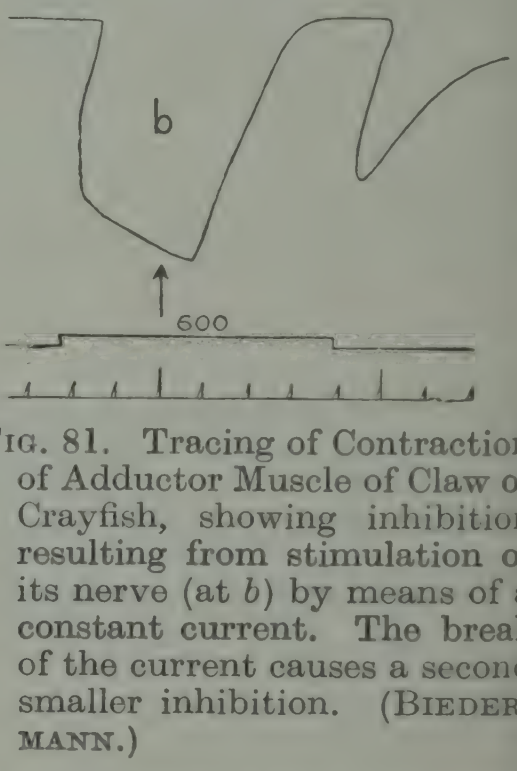


FIG. 81. Tracing of Contraction of Adductor Muscle of Claw of Crayfish, showing inhibition resulting from stimulation of its nerve (at *b*) by means of a constant current. The break of the current causes a second smaller inhibition. (BIEDERMANN.)

At the one extreme, smooth muscle preserves many of the properties of primitive protoplasm; it is distinguished from voluntary muscle not only by its histological structure, in its chemistry, its innervation, and its physical properties, but especially by the phenomena of independent tonus and rhythmic contraction. The nerve supply is often in the form of a nerve net, but is not indispensable for contraction, though capable of causing it and of modifying it. The plain muscle of hollow viscera is of this type. The individual fibres are fusiform cells with an oval centrally placed nucleus; the cytoplasm in the living cell appears to be structureless, and there is a delicate cell wall. The cells are probably connected together by fine intercellular bridges to form a syncytium, by which in all probability the excitatory process can be transmitted from cell to cell, though under conditions which are at present unknown.

---

* The *retractor penis*, found in the dog, cat, horse, hedgehog (but not in rabbit or man), is a thin band of longitudinally arranged unstriated muscle, which is inserted at the attachment of the prepuce and is continued backwards in a sheath of connective tissue to the bulb, where it divides into two slips, which pass on either side of the anus. It is innervated from two sources, the *motor* fibres being derived from the lumbar sympathetic and running to the muscle in the internal pudic nerve, while the *inhibitory* fibres run in the pelvic visceral nerves (*nervi erigentes*) and are derived from the second and third sacral nerve roots. (LANGLEY and ANDERSON. *J. Physiol.*, 1895, **19**, 88.)

The other extreme type, though similar in histological appearance, is under more precise and localised control, through the action of nerves. It does not show, or shows less readily, rhythmic spontaneous contractions. The plain muscle of blood vessels, the *arrectores pilorum*, the nictitating membrane, the intrinsic eye muscles, and the retractor penis belong to this type.

In most smooth muscle, there is a double nerve supply, one variety being motor, and causing contraction, the other kind inhibitory, and causing cessation of a previous contraction. All these fibres belong to the autonomic system, *i.e.* they arise from ganglion cells outside the central nervous system, and only receive impulses from the central nervous system through the intermediation of these ganglia. Usually one supply is sympathetic, and the other parasympathetic, *e.g.* to the intestinal muscle, sympathetic and vagus: their ultimate ramifications in the muscle form a plexus, and are non-medullated, and not infrequently the ganglion cells from which they arise are also contained in the nerve net near to or among the plain muscle cells which they supply. It is believed that both sources of supply are present in the nerve plexus, *e.g.* both vagus and sympathetic fibres are present in the plexus of Auerbach in the intestine.

The physiological significance of the double supply is evident from the fact that, even when deprived of all nerves, plain muscle may exhibit either sustained spontaneous contractions, known as tonic, or else intermittent rhythmic contractions, so that control of both contraction and relaxation is feasible (Fig. 80). In some invertebrates many voluntary striated muscles also possess a double nerve supply, one causing contraction of the relaxed muscle, and one causing relaxation of the contracted muscle (Fig. 81).

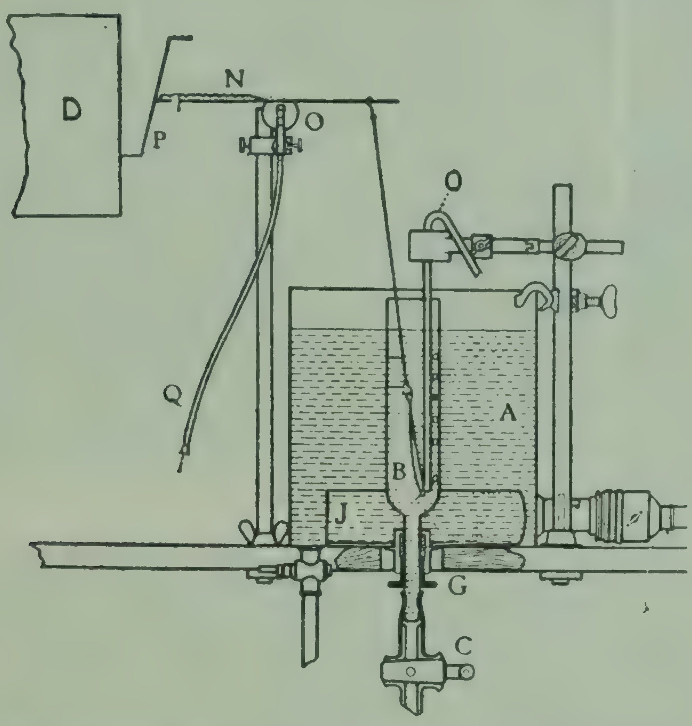


FIG. 82. Thermostat Bath for Investigation of Plain Muscle. A = copper water-bath, kept warm by the lamp J. B = Glass vessel filled with warm saline which can be changed by the tap C. O = the oxygen tube, at the bottom of which is a platinum pin with which the tissue is transfixed. N = a frontal writing lever, to which the tissue is attached by a hair. O, Q = a small brake for fixing the lever while the fluid is changed in B. D = recording drum. (BURN and DALE.)

Physiological Properties. In studying the responses of plain muscle it is not customary, as with amphibian skeletal muscle, to use the moist tissue suspended in air, because under these conditions the tissue behaves irregularly, but, instead, to suspend it in a bath of saline solution of suitable composition, *e.g.* Ringer's fluid, which is kept at the proper temperature and oxygenated with a constant stream of oxygen bubbles. A suitable arrangement is shown in Fig. 82.

TONUS. Plain muscle, isolated in this manner, is very readily extensible, so that when a small weight is applied to it the tissue may extend to several times its former length. The extension is, however, usually very slow, as though the cell contents were extremely viscous; also, the amount of lengthening finally produced by a weight will depend on the degree of *tonus* which the preparation has, and this varies with the source.* Thus the cat's uterus is highly tonic and requires a much larger weight to stretch it than does a guinea-pig's uterus of the same size. Many drugs have the property of affecting the

---

* WINTON. *J. Physiol.*, 1930, **69**, 393.

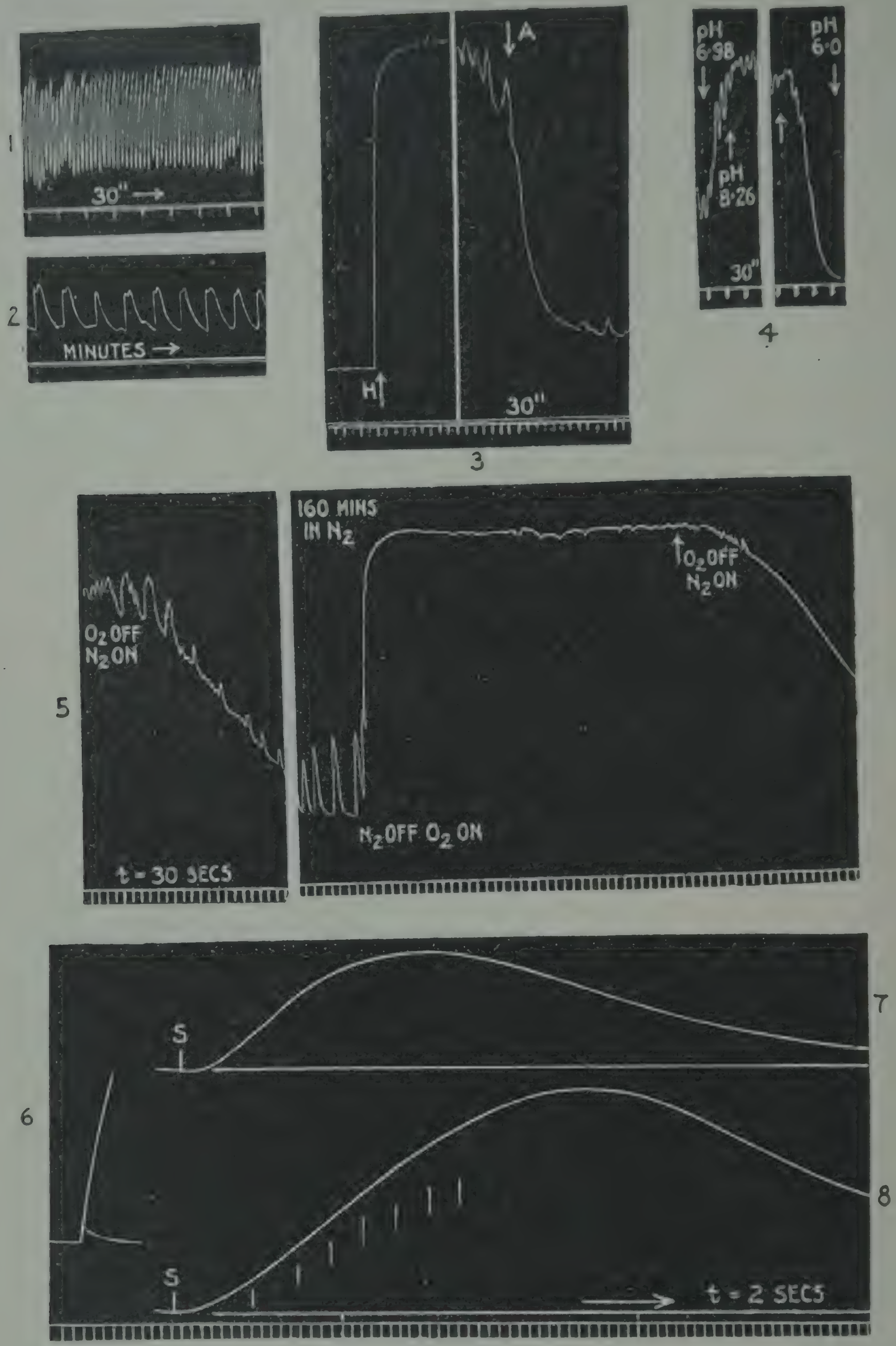


FIG. 83. 1. Isolated jejunum of rabbit, showing rhythmic contractions. Time tracing = 30 seconds.

2. Slow rhythmic contractions of guinea-pig uterus. (Time = minutes.)

3. Guinea-pig uterus. At H, histamine 1 in 1 million; at A, adrenaline, 1 in 4 million. Time = 30 seconds.

4. Small intestine of cat, showing contraction on raising, and relaxation on lowering the *p*H of the fluid in the bath. Time = 30 seconds; *p*H measurements taken at arrow. (EVANS and UNDERHILL.)

5. Uterus of cat, showing effect of oxygen lack caused by replacing oxygen bubbles by nitrogen. The second portion is taken after about two and a half hours in nitrogen.

6. On left, contraction of frog's gastrocnemius to compare with contraction of frog's stomach (7) on same slow drum. Point of stimulation = S. 8. Summation of contraction in frog's stomach. The first stimulus is followed by seven others at the points shown.

tonus of plain muscle, perhaps by causing polarisation changes at the surface of the cells. Thus adrenaline will cause the uterus to relax (if tonic), while acetyl choline even in high dilutions (*e.g.* 1 part in 10 million of saline) or histamine causes intense tonic contraction (Fig. 83, 3). Adrenaline in many other respects reproduces the action of the sympathetic nerves, and is therefore called *sympathicomimetic*, while acetyl choline reproduces the action of the parasympathetic nerve supply.

The maintenance of the sustained tonus is apparently not associated with increased expenditure of energy,[*] for the tissue, while in this condition, uses no more oxygen than when in the relaxed state, even though it may be exerting quite considerable tension; consequently no signs of fatigue are manifested.

Thus in the bivalve mollusc Pecten, there is a large striated adductor muscle which can close the shell rapidly but then relaxes, and a smaller smooth muscle whose office it is to hold the two valves together. The weight required to separate the valves, when this latter muscle is contracted, is many times that which will prevent the closure of the shell when open. Thus the smooth adductor muscle can support a weight which it cannot raise. This "catch action" is under the control of the central nervous system by two nerves from the cerebral ganglion. Stimulation of one of these causes contraction, and the muscle remains contracted even after the nerve is divided. Stimulation of the other nerve causes the muscle to relax. Apparently the action of these nerves is to allow the length of the muscle to alter, and it then remains at that length until the other nerve is stimulated. Somewhat similar phenomena may be observed in the smooth muscle of higher animals, and it may be that most smooth muscles are capable of two kinds of shortening, one followed by passive relaxation and the other by fixation at a diminished length without the expenditure of extra energy.

Although plain muscle can execute contractions for some time after being completely deprived of oxygen, and although, when in the tonic state, the muscle uses no more oxygen than when it is relaxed, yet the presence of oxygen is essential for the maintenance of tonus (Fig. 83, 5). It would seem as though oxygen, though not essential for the act of shortening, plays some part in maintaining the integrity of the cell structures which are concerned in fixing the tissue at the shorter length. (Garry.)[†]

Rhythmic Contractions. Many varieties of plain muscle when slightly stretched normally exhibit rhythmic contractions at fairly regular intervals, both in the body and when isolated. The strength and frequency of these rhythmic contractions vary very much according to conditions; in some types of plain muscle, *e.g.* that of the small intestine, they are regular, frequent and powerful (Fig. 83, 1), whereas in other cases they are slow (Fig. 83, 2), often irregular, and of variable amplitude, being occasionally negligible.

There is strong evidence that they originate in the plain muscle cells themselves, *i.e.* they are myogenic; for example, the cells in the chick amnion, which are devoid of nerves, show these rhythmic contractions[‡]; further, strips of the inner muscular coat of the small intestine, also free from nerve structures, were found[§] to exhibit rhythmic contractions.

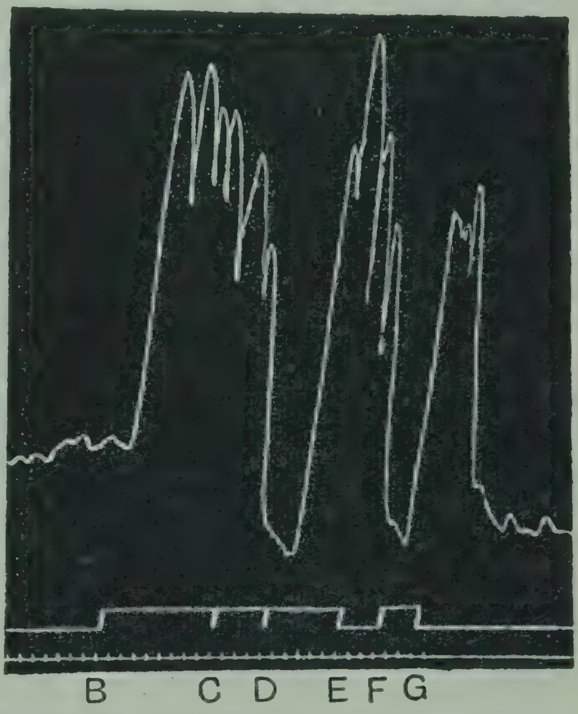


Fig. 84. Double-nerve preparation of cat stomach fundus. The vagus supply was stimulated (faradic) from B to E. At CD stimulation of sympathetic supply also, 12 shocks per sec., causing relaxation. The same stimuli to sympathetic alone at FG. (McSwiney and Robson.)

Stimulation. Smooth muscle responds to very various stimuli—electrical, mechanical, thermal, light, chemical, etc. As we should expect, however, from the sluggish response of this kind of contractile tissue (high chronaxie), the optimum rate of change of current which excites is very much slower than in the case of

[*] Lovatt Evans. *J. Physiol.*, 1923, **58**, 22.

[†] Garry. *J. Physiol.*, 1928, **66**, 235.

[‡] Ferguson. *Amer. J. Physiol.*, 1940, **131**, 524.

[§] Gunn and Underhill. *Quart. J. Exp. Physiol.*, 1914, **8**, 275.

striated muscle. Thus in many instances a single induction shock is powerless to excite contraction. A better stimulus is the make of a constant current.

Nerve-muscle preparations of mammalian plain muscle (stomach) with either or both of the nerves may be made.* Stimulation of the sympathetic supply with slow shocks usually caused contraction ; with more rapid stimuli relaxation. Stimulation of the vagus supply caused contraction. When sympathetic stimulation was applied during contraction due to vagal stimulation, the muscle relaxed, though the same stimulus to the sympathetic alone caused contraction (Fig. 84).

The contraction curve of plain muscle, though very variable, is much slower than that of voluntary muscle, and the phenomena of latent period, shortening, and relaxation can easily be followed with the eye. The latent period may be from 0·2 to one or two seconds, and the period of contraction from half a second to a few minutes, with proportional durations of relaxation (Fig. 83, 7) ; the stronger the stimulus the shorter the latent period and the higher the contraction.

As in voluntary muscle, if one stimulus follows another at a sufficiently short interval, we may have an increased liberation of energy due to repetition of the stimulus during the rise of the excitatory condition (*summation of contractions*) (Fig. 83, 8). As might be expected the refractory period is also longer in involuntary muscle.

In all involuntary muscle we may observe summation of the effects of stimuli, when each individual stimulus is insufficient to produce any excitation. Thus we may find a strength of induction shock which, applied singly, is just insufficient to evoke any response. If, however, the shocks are repeated at intervals of a second, it will be found that the first three or four stimuli are ineffective, and then the muscle enters into a contraction which increases with each succeeding stimulus until it has attained its maximum. There is thus *latent addition* of excitation. Each stimulus, in fact, alters the state of the contractile tissue and makes it more ready to respond to the next stimulus.

Latent addition is a general phenomenon of excitation, but appears to be of special importance in plain muscle. A plain muscle fibre may receive several nerve fibres and may only respond after the effects of several nerve impulses have been summated.†

It has been shown ‡ § that, as with other forms of muscle, the tension developed on contraction by plain muscle increases with increase of initial length, a very important property in view of the common occurrence of plain muscle in the walls of hollow viscera, where it is subjected to various degrees of extension.

Chemical Stimulation. The tissue is very sensitive to alterations in the hydrogen ion concentration ; within limits, lowering of H-ion causes contraction or tonus, with increased rate of rhythmic contractions, while‖ increase causes the reverse (Fig. 83, 4). Many drugs, such as eserine, salts of lead and barium, histamine, acetyl choline, pilocarpine, may cause contraction. As one would expect, however, from the greater independence of the smooth muscle, the action of these drugs varies from organ to organ, muscle fibres apparently histologically identical reacting diversely according to their origin. Some drugs, *e.g.* barium salts and perhaps pituitrin, appear to act directly on the muscle cells, while others may exert their action on the accompanying nerves or their endings. Adrenaline and acetyl choline in general evoke the same effects as stimulation of sympathetic or parasympathetic nerves respectively.

The question arises whether nervous stimulation may not produce its effects by the liberation of active chemical substances at the nerve terminations, or in the nerve network.

---

* McSwiney and Robson. *J. Physiol.*, 1929, **65**, 124 ; 1931, **71**, 194 ; **73**, 141.

† Eccles and Magladery. *J. Physiol.*, 1937, **90**, 31.

‡ Brocklehurst. *J. Physiol.*, 1926, **61**, 275.

§ Winton. *Ibid.*, p. 368.

‖ Lovatt Evans and Underhill. *J. Physiol.*, 1923, **58**, 1 ; McSwiney and Newton. *J. Physiol.*, 1927, **63**, 51.

Extracts of plain muscle certainly contain substances which can cause other plain muscle to contract,* and the active constituent in some extracts is probably acetyl choline.† It has been shown also that when plain muscle is made to contract by stimulation of nerve, substances are set free which can cause other plain muscle to contract. The action of these substances, like that of acetyl choline, is enhanced by eserine and counteracted by atropine.

Similar relations have been found by Finkleman ‡ and by Cannon and his associates § for sympathetic stimulation. When the sympathetic supply to almost any plain muscle is stimulated an adrenaline-like substance which Cannon has called "sympathin" is liberated and may enter the blood stream, in which case it causes acceleration of the denervated heart and other manifestations of sympathetic activity.

Mechanical Stimulation. Smooth muscle is very sensitive to mechanical stimuli, and may respond by a local or a general (propagated) contraction. One form of mechanical stimulation is that produced by tension. The effect of increasing the tension on smooth muscle may be twofold: in the first place extension and in the second excitation with increased contraction. These two effects are illustrated in the bladder. If this has all its connections with the central nervous system severed, it is when empty in a state of tonic contraction. If fluid be injected into it rapidly, there is a great rise of pressure in its cavity due to the forcible distension. If, however, the fluid be injected slowly, the bladder muscle relaxes to make room for it, so that a considerable amount of fluid may be accommodated in the bladder without any great rise of pressure. If the injection of fluid be continued, the walls begin to be stretched passively, and this increased tension acts as a stimulus causing rhythmic contractions of the whole bladder.

A peculiar property of plain muscle is that a stimulus (*e.g.* electrical or mechanical) which, when applied to the relaxed muscle, causes a contraction, will, when applied to the tonically contracted muscle, often provoke a rapid relaxation, and *vice versa*. The same even holds true in some instances for the effects of stimuli applied to the nerves supplying plain muscle. This property is perhaps responsible for the occurrence of rhythmic contractions when the muscle is given the constant stimulus of a slight stretch; this alternately excites the extended muscle to contract and the contracted muscle to relax.

ELECTRICAL POTENTIAL CHANGES. The visceral plain muscle, which is practically a syncytium, shows spontaneous contractions, and these are preceded by propagated potential changes, which are usually proportional to the power of the contraction, *i.e.* to the number of cells involved simultaneously. In tonic contractions there may be no propagated potentials, because the fibres are contracting asynchronously.

When the tissue responds to nerve impulses, the potential changes are complex and, owing to the smallness of the units, difficult to interpret. The nictitating membrane of the cat ‖ responds by contraction when the cervical sympathetic is stimulated. A volley too small to give a contraction may nevertheless yield an electrical response, the N-wave, of units subliminally excited, and while this lasts spatial facilitation can be obtained, as is the case with striated muscle and with nerve. A stronger stimulus, which means that stimuli from more nerve fibres impinge on each of the plain muscle cells, gives, after a shorter latent period (about 10 msec.) a complex of up to three negative waves, N, A, B, followed by a positive wave (Fig. 85). The contraction begins after the potential waves; the latent period of contraction is shorter the more powerful the stimulus to the nerve, but is usually about 150 msec. The A and B waves are diphasic, propagated, and are followed by a refractory period of about 50 msec. They are due to all-or-none response of a number of units. The A wave is not seen in weak contractions, but in strong ones tends to replace the B one, which has a longer latent period.

When a volley is increased beyond a certain point, so that a contraction of some power results, there is a series of repetitive waves (R in Fig. 85) following the initial complex and at intervals of $\frac{1}{3}$ to $1\frac{1}{2}$ seconds. Since the injection of adrenaline gives rise to similar repetitive waves, it is concluded that the R waves following a volley of nerve impulses are likewise due to the action of an adrenaline-like excitant liberated at the nerve terminals.

A second volley of wave impulses causes summation of contractions, and if it falls after the refractory period of the muscle this is easily understood. But it also may cause summation when it follows the first stimulus by a much shorter interval; the summation in this case is accompanied by R waves and is explained as due to the action of the second volley liberating adrenaline, although not directly exciting the muscle cells.

---

* LE HEUX. *Pflüger's Arch.*, 1918, **173,** 8; 1920, **179,** 177; 1921, **190,** 280.

† DALE and DUDLEY. *J. Physiol.*, 1929, **68,** 97.

‡ FINKLEMAN. *J. Physiol.*, 1930, **70,** 145.

§ CANNON and BACQ. *Amer. J. Physiol.*, 1931, **96,** 392; ROSENBLEUTH and CANNON. *Amer. J. Physiol.*, 1932, **99,** 398.

‖ ECCLES and MAGLADERY. *J. Physiol.*, 1937, **90,** 31, 68.

Propagation of the Excitatory State, or Wave of Contraction. In voluntary muscle there is no propagation from muscle fibre to muscle fibre, a contraction of the whole muscle being possible only when there is simultaneous excitation of all its fibres. This isolation of the excitatory state is not so evident in plain muscle. Frequently a stimulus applied to any part of a sheet of plain fibres may travel all over the sheet just as if it were a single fibre. Indeed there is protoplasmic continuity by means of fine bridge-like processes between adjacent cells in some varieties of plain muscle, such as in those occurring in the hollow viscera, and such tissues could act as syncytia.

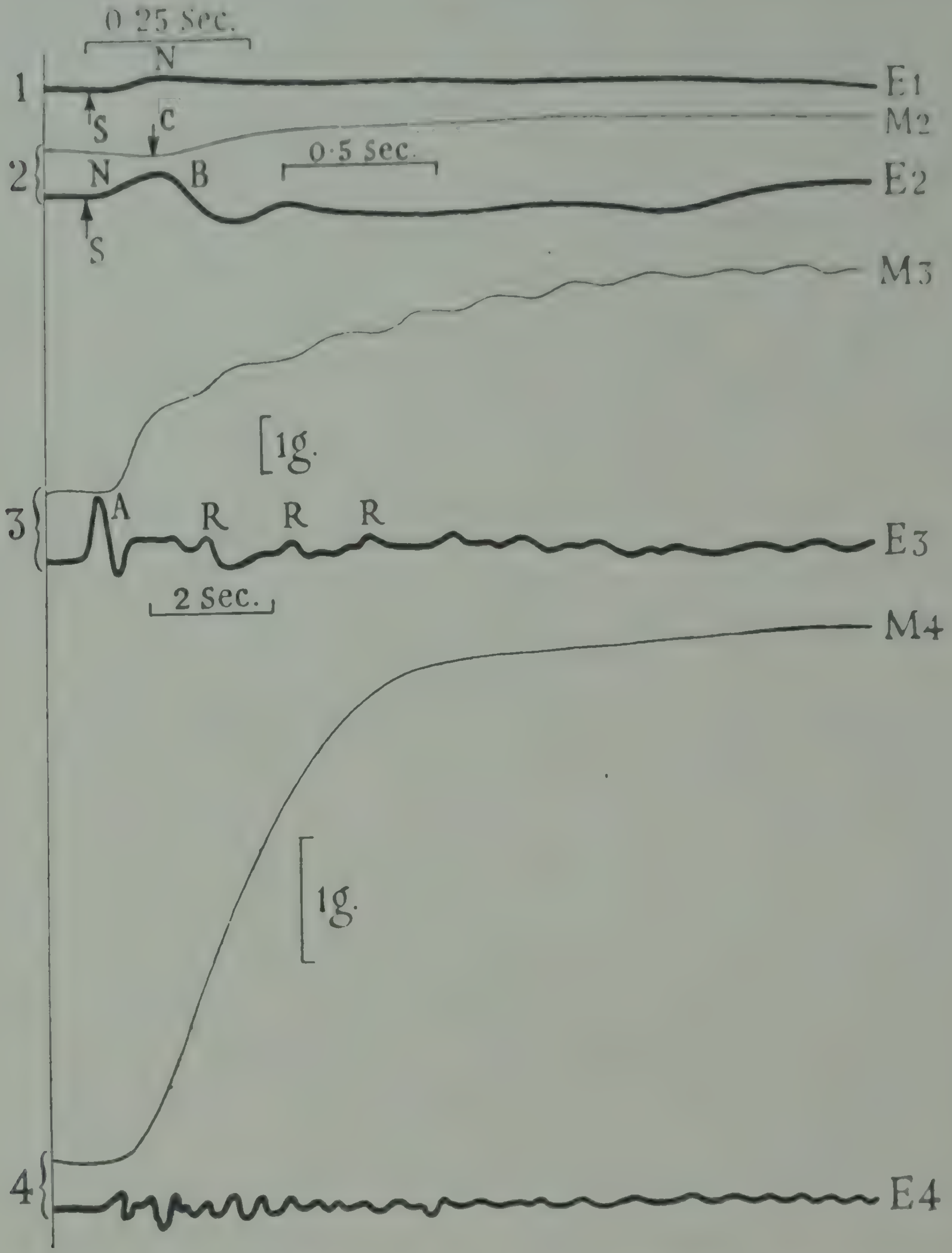


Fig. 85. Electrical potentials E (upward = negativity of attached border), and mechanical record of contraction, M (no contraction in 1), of nictitating membrane of cat resulting from stimulation of cervical sympathetic with stimuli of increasing strength from 1, to 3, and at 4 from injection of 50μg. of adrenaline into the circulation. S, stimulus ; C, commencement of contraction : N, A, B, initial potential changes : R, rhythmic waves. (Eccles and Magladery, *J. Physiol.*, 1937, **90**, 31, 68.)

It is claimed that each fibre of smooth muscle gives an all-or-none response to a stimulus.* Probably the contraction curve as commonly recorded represents the simultaneous or the successive action of the component cells, and a feeble contraction is probably due to the all-or-none response of only a few of them. It is certain, for instance, that feeble stimuli may produce only local contraction, especially in fatigued tissue and in those forms of plain muscle more directly under nerve control, and perhaps not possessing cell bridges. Thus, in the retractor penis of the dog, we get graded contractions from graded stimuli, and it is noticeable that with increasing strength of stimulus a greater extent of muscle enters into contraction.

It must be remembered that in most unstriated muscle the fibres are surrounded by a network of non-medullated nerve fibres. Some physiologists are inclined to ascribe to these fibres an important part in the propagation of the contraction wave.

---

* Fletcher. *J. Physiol.*, 1938, **91**, 172.

Influence of Temperature. Plain muscle is extremely susceptible to changes of temperature; as a rule, warming causes relaxation, while application of cold causes a tonic contraction. The condition of the muscle at any given time depends not only on its actual temperature, but also on the rapidity with which that temperature has been reached. Thus a rapid cooling of the retractor penis muscle of a dog from 35° to 25° may cause a contraction as extensive as would be produced by a slow cooling to 5° C. On warming plain muscle from 30° to 50° C. it lengthens gradually up to about 40°, and it may then undergo a marked heat contraction (varying in degree in different muscles) at about 50° C., which may pass off at a somewhat higher temperature. It is first paralysed and then killed somewhere between 40° and 50° C.* It seems very doubtful whether any true rigor mortis occurs in smooth muscle. The hard contracted appearance of the smooth muscle in a recently dead animal is chiefly conditioned by the fall of temperature. On excising the muscle and warming it up to body temperature it may again relax and show signs of excitability two or three days after the death of the animal.

The Chemistry of Contraction in plain muscle, so far as it has been studied, is very similar to that in striated muscle.† The most striking differences in chemical composition between striated and plain muscle are that the latter contains much less soluble protein, only very small quantities of creatine (and phosphagen), smaller amounts of glycogen, less potassium and phosphate and more sodium than the former. In the mussel, *mytilus*, arginine phosphate, the phosphagen of invertebrates, is broken down after prolonged stimuli and restored by recovery. Glycogen disappears and lactic acid arises when the muscle contracts, or when it is asphyxiated or treated with arsenates. The action of caffeine is, however, different, for this retards lactic acid formation in plain muscle. During recovery from contraction lactic acid is removed and glycogen restored. The amounts of lactic acid produced and of glycogen removed in contraction are relatively small, as would be expected in view of the small amount of energy it yields.

For Reference

Fischer (1944). Vertebrate Smooth Muscle. *Physiol. Rev.*, **24**, 467.
Lovatt Evans (1926). The Physiology of Plain Muscle. *Physiol. Rev.*, **6**, 358.

AMŒBOID MOVEMENT

Amœboid movement is seen in the unicellular organisms such as the amœba and in the polymorphonuclear blood corpuscles. It can occur only within certain limits of temperature (about 0° to 40° C.); within these limits it is the more active the higher the temperature. At about 45° the cell goes into a condition resembling heat rigor.

The fluid in which the corpuscles are suspended is of great importance. Distilled water, almost all salts, acids and alkalies, if strong enough, stop the action and kill the cell.

The movements are also arrested by CO_2 or by absence of oxygen.

Artificial excitation, whether electrical, chemical or thermal, causes withdrawal of the pseudopodia and contraction of the corpuscle, which therefore assumes the spherical form. Electrical stimulation also causes (*e.g.* in *Amœba*) a sudden gelation of the cytoplasm, which sets quite suddenly; ‡ after the stimulus is over, the fluid state returns.

For Reference

J. Gray (1931). Experimental Cytology. (Cambridge.)

CILIARY MOVEMENT

Cilia are met with in man in nearly the whole of the respiratory passages and the cavities opening into them, in the generative organs, in the Fallopian tubes of the female and the epididymis of the male, and on the ependyma of the central canal of the spinal cord and its continuation into the cerebral ventricles.

The cilia are delicate tapering filaments which project from the hyaline border of the epithelial cells. There may be twenty or thirty to each cell. The hyaline border is really made up of the enlarged basal portions of the cilia and seems to consist of a row of multiple centrioles.

The trachea of a mammal slit open along the back and pinned out under warm oxygenated Ringer's solution forms a suitable object for demonstrating ciliary action. If a drop of graphite suspension be placed on the mucous membrane, it will be seen to become entangled in mucus and pushed towards the upper end of the trachea at the rate of 1 to 2 cm. per minute. By this means any dust particles trapped on the ciliated surfaces of the respiratory tract are swept upwards to the mouth.

In action the cilia bend suddenly down into a hook or sickle form, and then return slowly to the erect position. This movement is repeated many (twelve to twenty) times a second, and thus serves to move forward mucus, dust or an ovum, as the case may be.

---

* Lovatt Evans. *J. Physiol.*, 1921, **54**, Proc. cxxxix.
† Lovatt Evans. *Biochem. J.*, 1925, **19**, 1115; 1926, **20**, 893.
‡ Bayliss. *Proc. Roy. Soc.*, 1919, **91B**, 196.

During the rapid, or effective, phase the cilium appears to be rigid, while during the slower return phase it is limp. No structure is demonstrable within the cilium to account for its power of movement. The minuteness of the object has up to now prevented us from deciding whether the cilium is itself actively motile, but the somewhat similar flagellum of certain unicellular organisms does seem to generate at least part of its own energy, irrespective of the attached cell.* An attractive suggestion, due to Gray, is that the movement may be due to a rhythmic alteration in the distribution of water in the cilium ; for example, a transference of water to one side would cause that side to swell, and so, like paper wetted on one surface, to bend over. The movement is apparently dependent upon connection of the cilia with their basal granules, since when the ciliated hem is stripped away, movement ceases at once.

There seems to be a functional connection between all the cells of a ciliated epithelial surface. The action is passed from cell to cell and from cilium to cilium, in succession. When looked at from above, a series of waves (*metachronal waves*) are seen to pass in regular sequence, over the ciliated surface, just as, when the wind blows, waves of bending pass over a field of corn.

Ciliary movement is not dependent upon nervous action, though in some instances it can be accelerated or inhibited through the action of nerves. Thus Seo claims that the rate of ciliary movement in the palate of the frog is accelerated by stimulation of appropriate nerves which are distributed to the palate and form a nerve-net. By the use of a " nerve-cilia " preparation consisting of the palate and palatine nerve, he finds that the velocity with which red blood cells can be swept along the ciliated surface is accelerated 7–20 times by stimulation of the nerve. This acceleration is due to more rapid metachronal waves, greater speed of effective phase and greater amplitude of movement of cilia. Normally he claims that the regulation of ciliary action is a reflex confined to the same side of the palate as the stimuli.

The conditions of ciliary action are the same as those for amœboid movement of naked cells. Acids or excess of calcium ions slow, and alkalies or potassium ions hasten, the rate of action. Rise of temperature, within limits, hastens, and cold slows, the rate, but over most of these ranges the amplitude of the movement is constant. Oxygen is utilised at a rate proportional to the rate of ciliary action, and though ciliary action will persist for a time in its absence, the movement ultimately fails if it is withheld (Gray).

For Reference

Gray (1928). Ciliary Movement. (Cambridge.)
Seo (1931). *Japanese Journ. Med. Sci.* III. Biophysics, **2**, 47.

---

* Lowndes. *Proc. Zool. Soc.*, 1941, **A. 111,** 111.

CHAPTER XII

NERVE FIBRES

THE STRUCTURE OF NERVE FIBRES

A STIMULATION of the nerve of a nerve-muscle preparation is followed after a very short interval by a contraction of the muscle. This observation illustrates the two functions of nerve fibres, excitability and conductivity—that is to say, a suitable stimulus can set up in any part of the nerve a disturbance which is propagated along the nerve without any visible effects occurring in it; and it is not until this *nervous impulse* has reached the muscle that a visible effect is observed in the shape of a contraction. In the animal body a direct excitation of the nerve fibre in its course never takes place normally. The only function

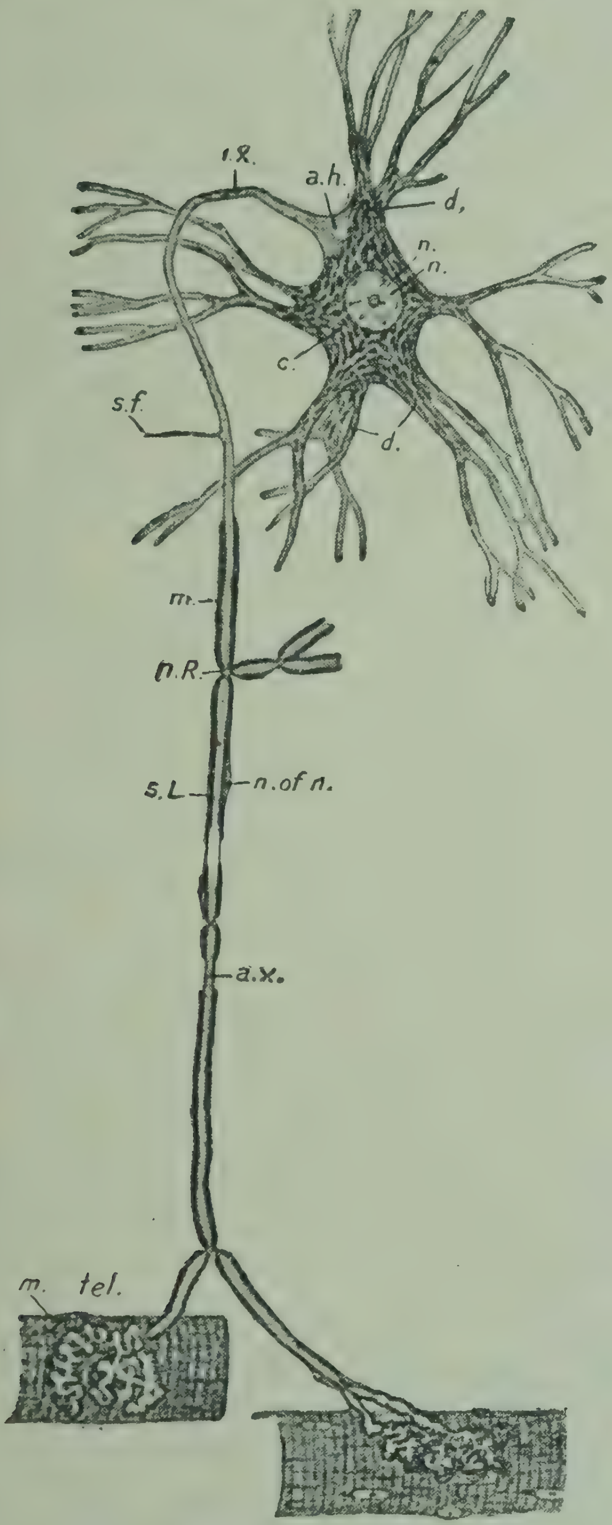


FIG. 86. Diagram of a Motor Nerve Cell with its Nerve Fibre and two motor-endplates. (After BARKER.) *a.h.* axon hillock; *d*, dendrites; *a.x*, axis cylinder; *m*, medullary sheath; *n.R*, node of Ranvier.

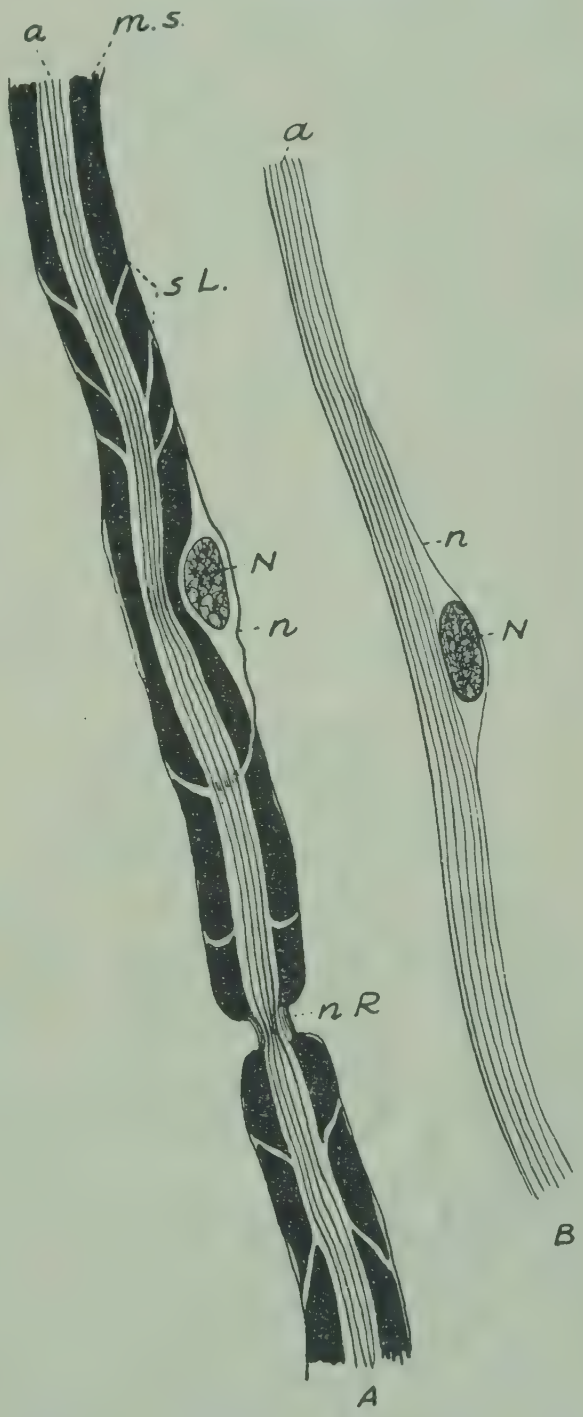


FIG. 87. A, portion of a medullated nerve fibre. *a*, the axis-cylinder, containing neurofibrils *m.s.*, the medullary sheath; *n*, the neurolemma; N, a nucleus; *n*.R, a node of Ranvier; *s*.L, segments of Lantermann (artefacts); B, portion of a non-medullated fibre. Lettering as for A.

the nerve fibre has to perform is that of conducting impulses, either afferent ones from the sense organs at the periphery to the central nervous system, or efferent impulses from this to the muscles or other effector organs. Hence it is essential that there should be continuity along the whole length of the fibre. Damage to any part, such as by crushing or heat, or cooling, causes a block to the passage of an impulse.

A nerve fibre is essentially a long process of a nerve cell (Fig. 86). The cell is in the higher animals usually withdrawn from the surface into a special collection of cells such as the posterior root ganglion, or may be one of the mass of cells and fibres making up the central nervous system. All nerves are alike in possessing as their conducting part the continuous strand of protoplasm produced from the nerve cell and known as the *axon* or *axis cylinder*, carried in a delicate tubular sheath, the *axolemma*. In the fresh condition the axon appears to be a homogeneous fluid, though after fixation special staining methods bring out the presence of *fibrillæ* or *neurofibrils*, embedded in a clear material, the *axoplasm* (Fig. 87). There is a difference of opinion as to the existence of neurofibrils in the living nerve fibre. In the giant nerve fibres of Cephalopods, which may reach 1 mm. in diameter, the contents are fluid and run out when the fibre is cut across.* But micro-dissection of living nerve fibres in other invertebrates and also in vertebrates shows the contents of the axis cylinder to be gelatinous and to contain neurofibrils as tough threads.† It is conceivable that gelation results from the manipulations. Claims have been made that there is normally a continuous flow of axoplasm from the cell body along the axon.

The electron microscope shows ‡ a new and characteristic structure consisting of long, thin fibres (200–1,000 A. diam.) with dense edges and a less dense axis, and which show periodic cross-bandings; they appear to be tubular and are called "neurotubules."

In the course of growth the nerves develop certain histological differences, which appear to bear some relation to the character of their parent cell. Thus many, or, all the fibres which are given off from and which enter the central nervous system, *i.e.* the brain and spinal cord, belong to the class known as *medullated*. In this type the axon is surrounded with a layer of material known as *myelin*, forming the *medullary sheath*. This sheath consists of a fatty material composed largely of lecithin, and staining black with osmic acid. Study of the optical properties of the myelin shows that the lipoid molecules are oriented radially to the axon, in concentric shells two molecules thick, and with protein between the successive double layers. The myelin is supported in the interstices of a network formed of a horny substance known as *neurokeratin*. The medullary sheath is surrounded by a structureless membrane, the primitive sheath or *neurolemma*. At regular intervals a break occurs in the medullary sheath, the neurolemma coming in close contact with the axis cylinder which is constricted at that place. This break is the *node of Ranvier*, the intervening portions of medullated nerve being the *internodes*. In each internode, lying closely under the neurolemma, is an oval nucleus. The medullated nerve fibres vary considerably in diameter, the largest fibres being distributed to the muscles and skin, the smallest carrying impulses from the central nervous system to the viscera (Fig. 88). The latter all come to an end in some collection of ganglion cells of the sympathetic chain or peripheral ganglia, the impulses being carried on to their destination by a fresh relay of non-medullated nerve fibres.

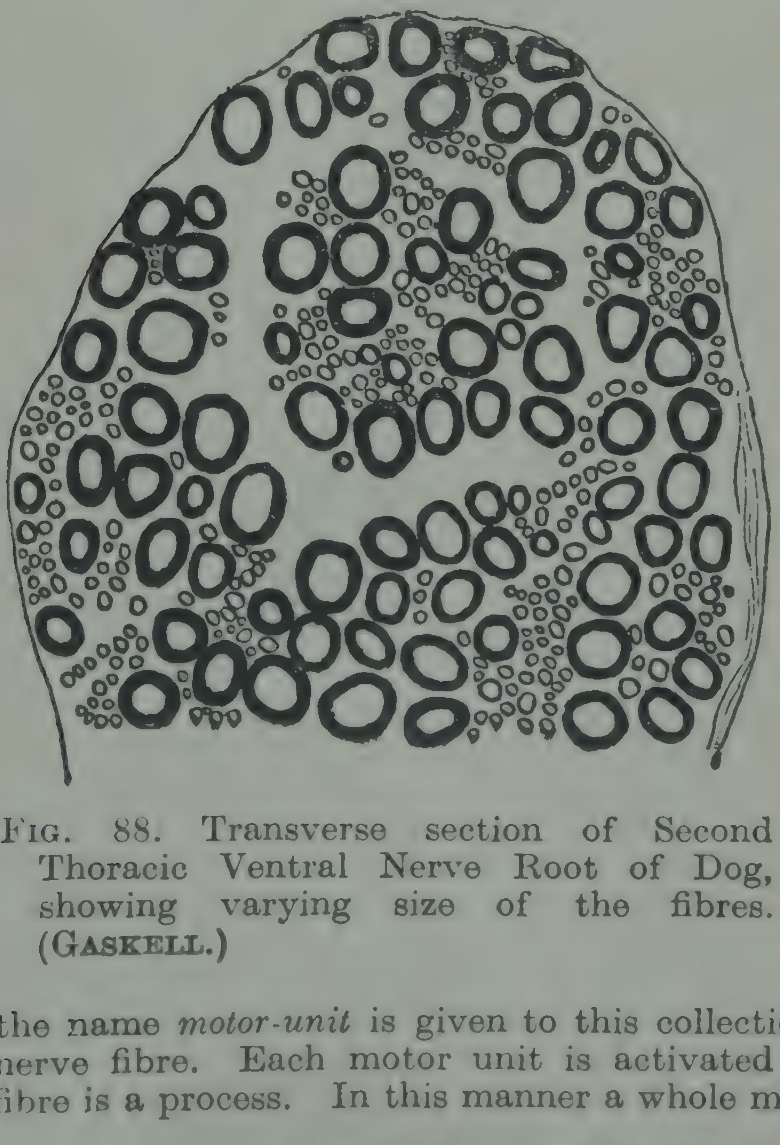


FIG. 88. Transverse section of Second Thoracic Ventral Nerve Root of Dog, showing varying size of the fibres. (GASKELL.)

The medullated nerves are bound together by connective tissue (*endoneurium*) into small bundles, which are again united by tougher connective tissue into larger nerve trunks. The fibres as a rule branch only when in close proximity to their destination, and then the branching always occurs at a node of Ranvier. The number of branches may be large; thus each motor fibre which arises from a motor cell may supply as many as 150 muscle fibres, and the name *motor-unit* is given to this collection of muscle fibres, together with its motor nerve fibre. Each motor unit is activated by one nerve cell, from which the motor fibre is a process. In this manner a whole muscle may be under the control of relatively few motor nerve cells and motor nerve fibres—*e.g.* the soleus of the cat has 30,000 muscle fibres grouped into 230 motor-units, *i.e.* its motor nerve contains only 230 fibres.

---

* YOUNG. *J. Physiol.*, 1935, **83**, 27P.

† DE RENYI. *J. Comp. Neurol.*, 1929, **48**, 441.

‡ DE ROBERTS and SCHMITT. *J. Cell. and Comp. Physiol.*, 1948, **31**, 1.

The blood supply to nerve trunks is small, but abundant for their needs, and indispensable for their functions.*

The *non-medullated fibres* (Fig. 87B) differ from the medullated simply in the absence of a medullary sheath. They possess a neurolemma, under which we find nuclei lying closely on the side of the fibre and bulging out the sheath. In their ultimate ramifications they tend to form close networks or plexuses and appear to lose the last traces of a sheath.

As to the functions of the myelin sheath in the medullated nerve fibre, very little is known. It does not make its appearance until the axis cylinder is formed, and is apparently derived from a series of cells which grow out from the spongioblasts of the central nervous system and form a chain surrounding the out-growing axons. In the regeneration of a nerve fibre after section, the myelin sheath appears later than the axon in the peripheral part of the nerve. It has been supposed to act as a sort of insulator ensuring isolated conduction within any given nerve fibre.

The Degeneration of Nerve Fibres. When an axon is severed, either by actual division, or functionally by localised damage by chemical or physical means, and so cut off from the cell upon which its life depends, the part so cut off undergoes rapid degenerative changes. For a few millimetres on either side of the cut there are acute changes known as *traumatic degeneration.* But the peripheral portion also shows, simultaneously along its whole length,† the changes of secondary or *Wallerian degeneration*, which commence in the first hour and progress to complete degeneration in a few days ; loss of excitability to stimuli is complete after three days. The central portion of the fibre does not undergo secondary degeneration, though it may atrophy later on (retrograde degeneration) as may the cell body, if no fibre regeneration takes place.

The structural changes during degeneration are initiated by the appearance of swellings and fusion in the neurofibrils and shortly afterwards the axis cylinder breaks up irregularly into short lengths. The myelin sheath, in medullated nerves, breaks up into irregular droplets and undergoes chemical changes, among which is the liberation of highly unsaturated fatty acids. In consequence of this the degenerating medullated nerve fibres are stained black when treated with Marchi's method, a fact of great importance for mapping out the tracts of the central nervous system. Meanwhile the cells of the neurolemma rapidly proliferate. Lastly, the proliferated neurolemma cells, together with histiocytes of the surrounding connective tissue, digest and absorb the remnants of the axis cylinder and myelin sheath, so that all that ultimately remains of the fibre is a thin thread consisting of modified neurolemma cells. The specialised terminal nervous structures in motor end-plates, sensory endings, etc., also undergo degeneration, but the protoplasm of the end-plates persists for months.

Regeneration of Nerve Fibres. Some time after section and degeneration, a process of regeneration may occur in certain nerves. The process occurs by the formation of outgrowths of neurofibrils from the undegenerated stumps which remain connected to their cells. These outgrowths have bulbous ends, and sprout down, probably by some sort of chemiotactic process, into the neurolemmal remains of the degenerated fibre, growing at a rate of about 1·5 mm. to 4 per day.‡ At the same time, if there is a gap, the cells of the old neurolemma proliferate and grow up until they establish continuity with the neurolemma of the fibres in the normal stump. The presence of the old track, preserved by the neurolemma cells, appears to be indispensable for the proper guidance of the growing axons. Hence, if the gap is great, the bulbous ends of the axons grope about in the tissues, and end by forming a tangled mass, or neuroma. Because of the necessity for the neurolemma also, regeneration never takes place in those fibres of the nervous system that have no neurolemma, *e.g.*, in the white matter of the central nervous system, the optic nerve, or the sensory root of the Vth cranial nerve.

---

* Bentley and Schlapp. *J. Physiol.*, 1943, **102**, 62.

† Sugar. *J. Neurophysiol.*, 1938, **1**, 7.

‡ Seddon *et al.* *J. Physiol.*, 1943, **102**, 191 ; Young. *Physiol. Rev.*, 1942, **22**, 318.

After the establishment of contact of the growing ends of the fibres with the old neurolemma tracks, growth steadily proceeds along these until, the terminations having been reached and regeneration of specific end-structures effected, re-establishment of function becomes possible. The medullary sheath is reformed shortly after the regeneration of the axis cylinder. The regenerated fibres tend to be thinner than the original ones,* except after mere crush injury. The extent to which structural and functional recovery takes place, and also its speed, varies greatly with the conditions attending the section, and with its site. If the division is a purely functional one, as by local crushing, recovery is relatively complete and rapid, especially if the length to be regenerated is a short one. Moreover, the degree of recovery of function is considerable, because each new fibre has a fair chance of growing into its former track.† If the nerve is actually severed and the cut ends sutured together there will be less rapid and less complete recovery of function; it has been found that a fuller recovery ensues in this case if the two ends are not brought together for a week or two, so that preliminary degeneration is complete before reunion is encouraged. If a considerable gap is left between the two divided ends recovery is very slow and often but slight.

These incomplete recoveries are ascribed to the fact that the nerve fibres, when they grow, enter the track of fibres of different function.‡ Thus in a mixed nerve, a sensory axon may grow down into the track of a motor fibre, and no functional result would follow. Or a sensory fibre, say from skin, may grow into another skin sensory track of different function, and in that case abnormal sensations would arise, for the removal of which re-education over a considerable time would be needed.

The reunion of motor nerves of different muscles has found application in surgery. Thus the central end of the divided spinal accessory has been united to the distal end of a paralysed facial nerve, and after regeneration and re-education has given good results.

Langley and Anderson § also united many different nerves, *e.g.* the central end of the vagus with the peripheral end of the cervical sympathetic and showed that the vagus fibres regenerated up to the synapses of the superior cervical ganglion. As a result of their researches, moreover, it is clear that what are now called cholinergic nerves cannot be united to adrenergic nerves.

For Reference

Lee (1929). The Regeneration of Nervous Tissue. *Physiol. Rev.*, **9**, 575.
Young (1942). The Functional Repair of Nervous Tissue. *Physiol. Rev.*, **22**, 318.

THE EXCITATION OF NERVE FIBRES

Methods. Many different forms of stimuli, such as thermal, mechanical or chemical stimuli may be used to arouse the activity of nerve. Sudden warming of the nerve always gives rise to excitation. At about 45° C. the nerve dies. A nerve may be rapidly cooled without any excitation taking place.

A nerve may be excited mechanically by crushing or cutting, or by carefully graduated taps. The chemical stimuli most used are strong salt solutions, glycerol or weak acids.

None of these forms of stimuli can be adequately controlled either as to strength or duration, and all cause some damage. We therefore use electrical stimuli. For this purpose we may use either the make and break of a constant current, the induced current, the discharge of a condenser, or alternating currents. A valuable method is provided by an electronic device furnishing square waves of any desired voltage, duration or frequency.

---

* Gutmann and Sanders. *J. Physiol.*, 1943, **101**, 489.

† Sharpey Schafer. *Brain*, 1927, **50**, 538.

‡ Stopford. Sensation and the Sensory Pathway. London. Longmans, Green. 1930.

§ Langley and Anderson. *J. Physiol.*, 1904, **31**, 365.

The relation of current strength and duration to excitation in nerve fibres is determined by methods similar to those used for muscle, and strength-duration curves having similar significance can be plotted, as already discussed. A nerve cannot be excited by currents passed transversely across it, the threshold for stimulation being inversely proportional to the cosine of the angle between current and nerve (cos 90° = 0).

Criteria of Excitation. As evidence that excitation (and conduction) has taken place we have the effects produced (*e.g.* in the muscle in the case of a motor nerve) at the tissue supplied, and also the passage of a wave of electrical disturbance along the nerve.

THE ELECTRIC POTENTIAL. The basal facts of bioelectric potentials have been dealt with in Chapters VIII and X.

When a nerve is stimulated so as to cause excitation, a transitory state of negativity results at the excited point, and this state of negativity travels as a wave at a definite velocity in both directions along the fibres, to be followed by recovery processes which restore the tissue to the resting state. This travelling disturbance is the nervous impulse. It has been shown that the normal impulses travelling up sensory nerves are of similar nature to those which normally travel down the motor nerves.

It should be made clear to beginners that when we say that an electrical disturbance travels along the nerve, this does not mean that electricity travels along the nerve, any more than if we carry an accumulator across a room it means that electricity travels across it. An electric current can be conducted through a nerve cooled to 0°, but a nervous impulse cannot pass in such a cooled nerve.

Usually each stimulus causes a single nerve impulse, though very powerful ones may cause two or three in succession (repetitive response). Lack of calcium increases the tendency to repetitive response.*

In order to demonstrate the wave-like progression of the electrical change from an excited spot along the nerve, it is now customary to amplify the waves by thermionic valves and to record by means of a cathode ray oscillograph. If we lead off from an injured and an uninjured spot, the passage of each impulse past the uninjured spot will give rise to a monophasic variation ; if both spots are uninjured the variation will be diphasic (Fig. 48). The amplitude is about 25 to 50 millivolts in the fibre.

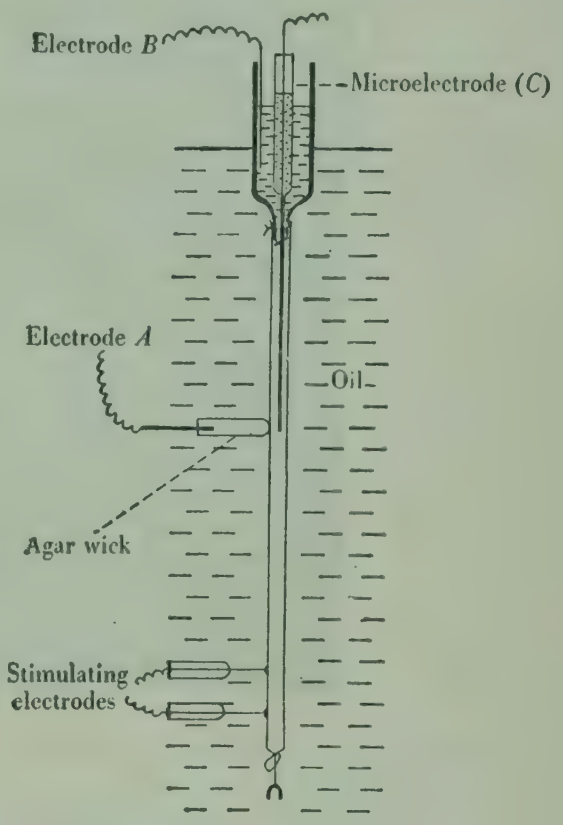


FIG. 89. Diagram of arrangements for recording action potentials in giant nerve fibre. Leads from A and C across membrane : from A and B the potential along membrane to cut end. (HODGKIN and HUXLEY. *J. Physiol.*, 1945, **104**, 176.)

What is important from the theoretical standpoint is the potential difference between the two sides of the plasma membrane, which is presumed to be near to, or identical with, the axolemma. With the usual arrangements for lead-off from the surface of the fibre, there is much short-circuiting, and so the real P.D. across the membrane is not recorded. These difficulties have been overcome by the use of giant non-medullated fibres of the squid, the axons of which may be 0·5 mm. in diameter. Into one of these fibres a cannula is tied and a microelectrode passed through it into the liquid axoplasm : this gives a lead-off from inside the membrane. The outer lead-off is opposite it, and short-circuits are reduced by immersing the fibre in oil (Fig. 89).

With such an arrangement, the resting potential across the membrane is seen to be about 40 millivolts; and this, as explained in Chapter VIII, is due to polarisation of the plasma membrane, the inside layer being negative. Now it has long been known that when an action potential passes along the fibre this resting potential is reduced, *i.e.* the membrane was thought to be depolarised. But in the squid fibre prepared as above, the membrane is not merely depolarised, but has its polarisation reversed, so that the inner surface is now positive to the outer (Fig. 90).

---

* KATZ. *J. Physiol.*, 1937, **88**, 239.

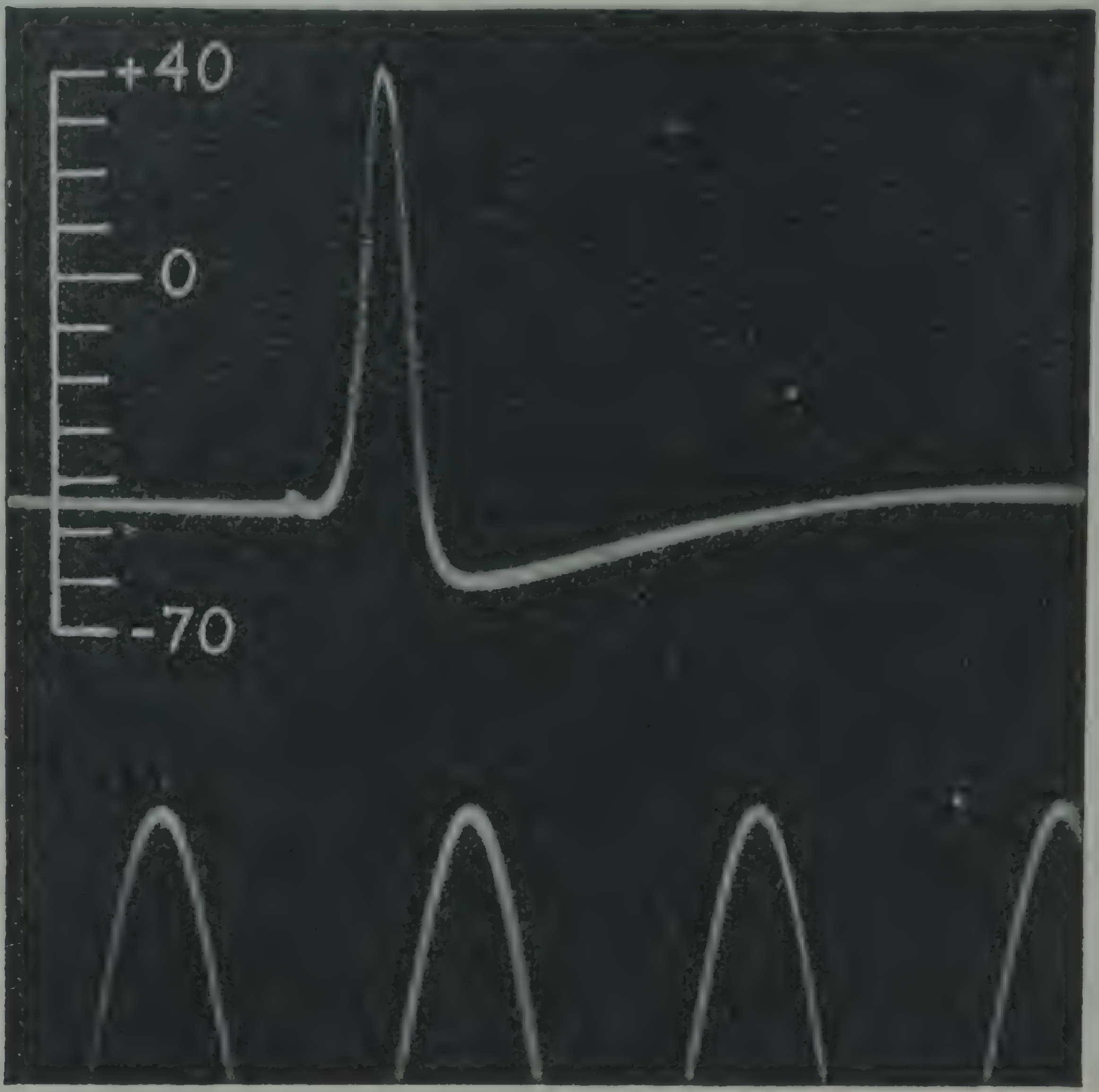


FIG. 90. Action potentials across membrane of single squid axon. The scale reads in mV. and the signs + and − refer to the charge inside the membrane. O = potential of surrounding sea water. On passage of the action potential the charge alters from —43 to + 40. Time = 500/sec. (HODGKIN and HUXLEY. *J. Physiol.*, 1945, **104**, 176.)

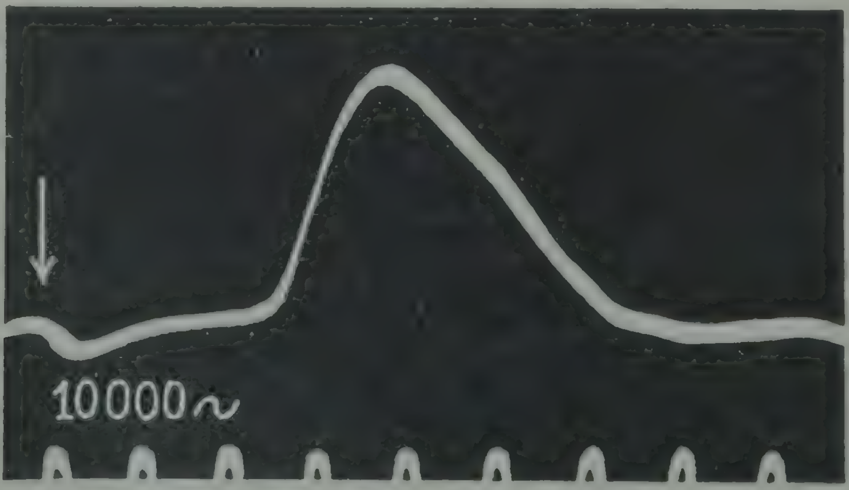


FIG. 91. Spike from single mammalian nerve fibre. Stimulus at arrow. The latent period is made up of true latent period at cathode plus conducting time to the lead. (GASSER. *J. Appl. Physics*, 1938, **9**, 88.)

This result is of great theoretical significance and has already been discussed in Chapter VIII.

The impulse progresses along the nerve at a speed depending upon the nature of the fibres, which fall into three main groups, A, B, and C. It travels at about 100 metres per second in the mammalian A-fibres, and lasts only an extremely short interval of time at each spot; thus the "rising phase" in mammalian A-fibres occupies 0·2 msec. and the total duration 0·4 msec. (Fig. 91), the figures for frog being 0·3 msec. and 0·9 msec. respectively.* Thus the *length* of the excitatory wave in mammalian A-fibres is about 40 mm.

This large monophasic wave is what is generally recorded as a means of recognising the passage of nervous impulses along nerve fibres under any given conditions of experiment, and is now referred to as a "*spike-potential*" to distinguish it from smaller and slower potential changes which modern refinements of technique have shown to follow it. These are respectively the *negative and positive after-potentials*.† (Fig. 92.) In mammalian fibres the positive after-potential cuts into the preceding negative after-potential. In the case of the A-fibres the spike lasts about 0·4 msec.; the negative after-potential continues the fall of negative potential, though more slowly, and may last 15 msec.; the positive after-potential is only about 0·2 per cent. of the magnitude of the spike, but lasts till 70 msec., and may then be succeeded by another negative after-potential. Tetanic stimuli (Fig. 93) or application of veratrine increase both after-potentials. For C fibres the spike lasts 2 msec., the negative after-potential up to 80 msec., and the positive after-potential for 1 to 2 seconds, and is large (1·5 per cent. of spike).

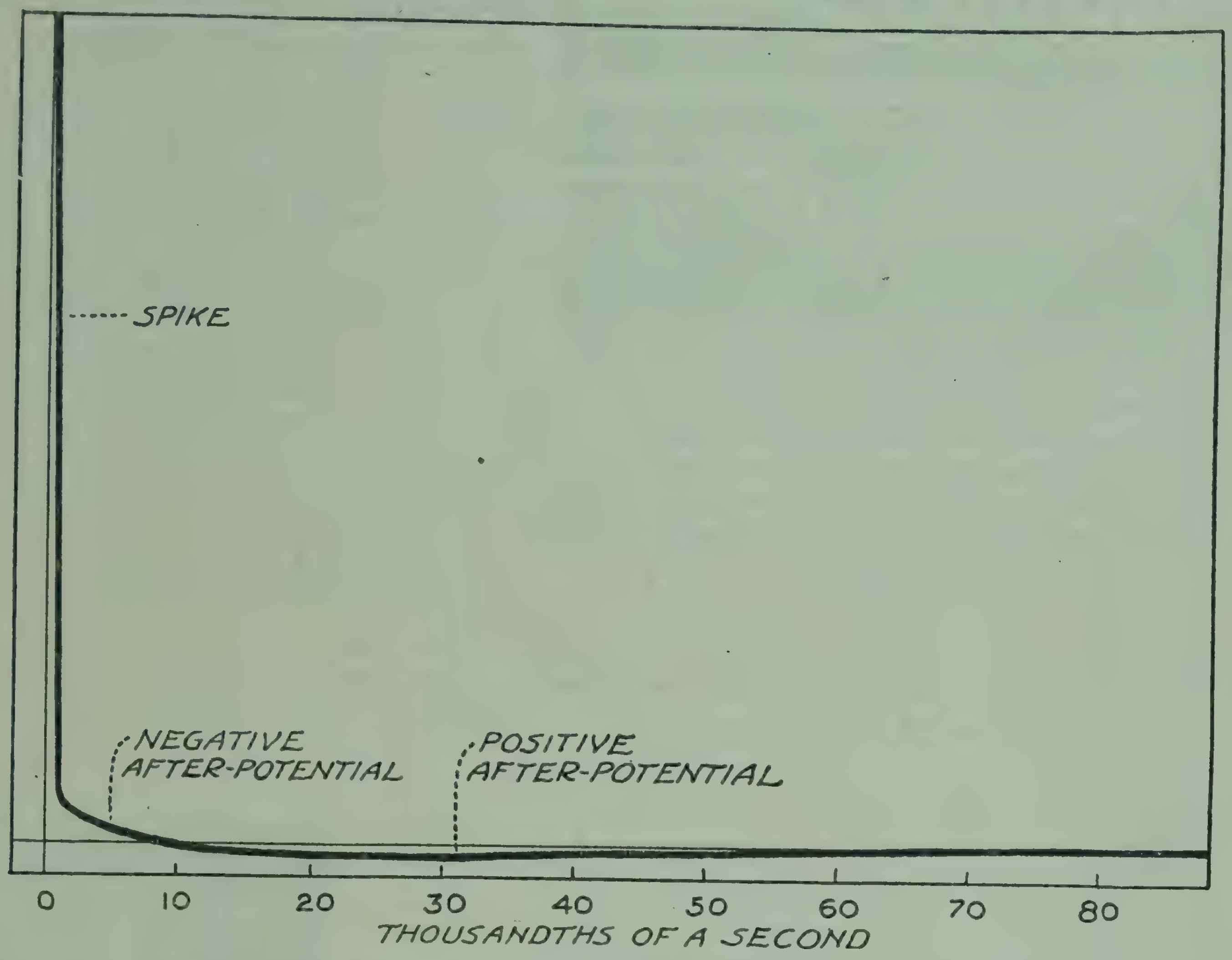


FIG. 92. Diagram of Electrical Potentials in Nerve following Stimulation at 0. The positive after-potential has been exaggerated. (After GASSER. Harvey Lect., 1937.)

The absolute refractory period in a nerve fibre occurs during the spike, the relative refractory phase is short and is followed by a supernormal phase lasting till the end of the negative after-potential, while the succeeding subnormal phase coincides with the positive after-potential.‡ It seems clear therefore that the after-potentials are an indication of the level of excitability, whereas the spike is the actual message carrier.

When by repeated stimulation the nerve is thrown into protracted activity, a study of the spike potentials indicates that the excitatory changes in the nerve fibre are *interrupted*, not continuous. We may compare the passage of such a stream of nerve impulses to the discharge of a stream of machine-gun bullets, rather than to the flow of a stream of water; "continuous" activity of the fibre being represented by a quick succession of these short waves (Fig. 93).

---

* GASSER and ERLANGER. *Amer. J. Physiol.*, 1927, **80**, 522.
† GASSER and GRUNDFEST. *Amer. J. Physiol.*, 1936, **117**, 113.
‡ AMBERSON. *Journ. Physiol.*, 1930, **69**, 60.

The maximum frequency at which the impulses can follow one another is determined by the length of the refractory period, since a second response cannot be obtained until the previous one has subsided. Since the absolute refractory period for mammalian A-fibres is about 0·4 msec., it follows that the maximum frequency is about 2,500 impulses per second ; the relative refractory phase, however, lasts for 3 msec., so that at frequencies greater than about 300 p. sec. all impulses after the first one will be slightly reduced (Fig. 94), and, incidentally, are also conducted along the fibre more slowly. When the frequency of stimulation of A-fibres is less than 300 p. sec., all the impulses are of full size, and this is the case whether the stimulus be strong or weak. In frog's nerve at 20° the maximum frequency is about 500 p. sec.

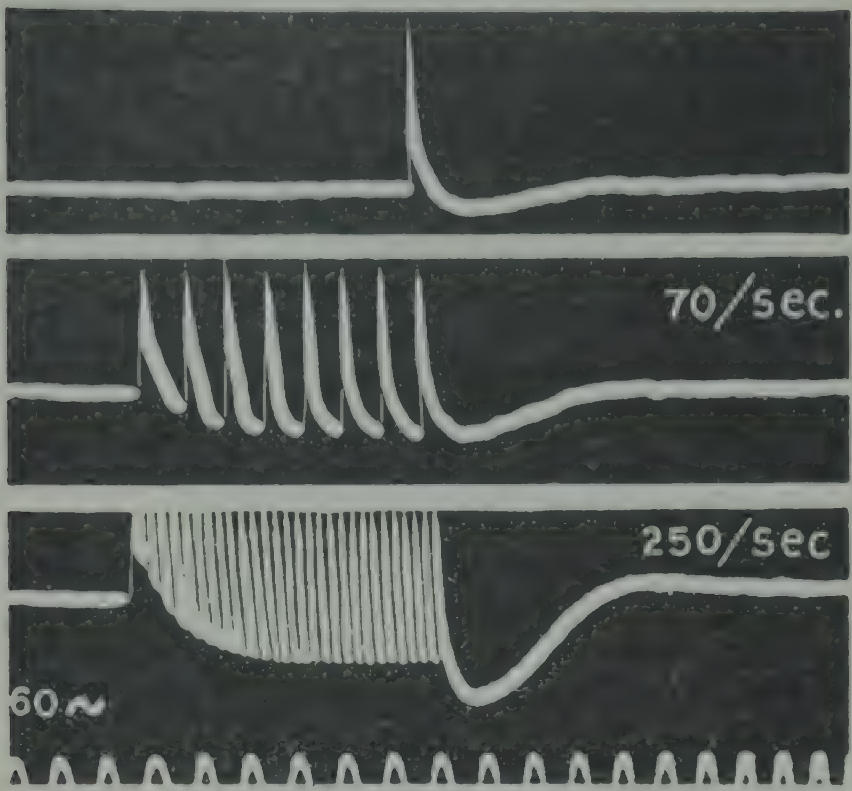


FIG. 93. Action and after-potentials in Phrenic Nerve, from single stimulus, from 8 stimuli at 70/sec., and from numerous stimuli at 250/sec. Note increase in positive after-potentials (dip below zero line) as frequency increases. (GASSER. *J. Appl. Physics*, 1938, **9**, 88.)

Volleys and Trains of Impulses. These two terms in frequent usage must now be explained. If a maximal stimulus be applied to a nerve at a given point, impulses are fired off simultaneously from that point along all the fibres. (A weaker stimulus will start off impulses along fewer fibres.) The result is in either case known as a *volley* of impulses, in fanciful comparison with the firing of a volley of rifle shots, the essence of the matter being that the impulses all start off at once, like scratch runners in a race. Actually there are, so to speak, two sets of runners, starting back to back and travelling in opposite directions, because the stimulus to each nerve fibre in its course initiates two such impulses in opposite directions. It will be seen that this is one way in which many impulses may be sent along a stimulated nerve, one along each fibre in each direction. It is clear that this gives a sort of *spatial summation*, corresponding to quantal summation in a muscle.

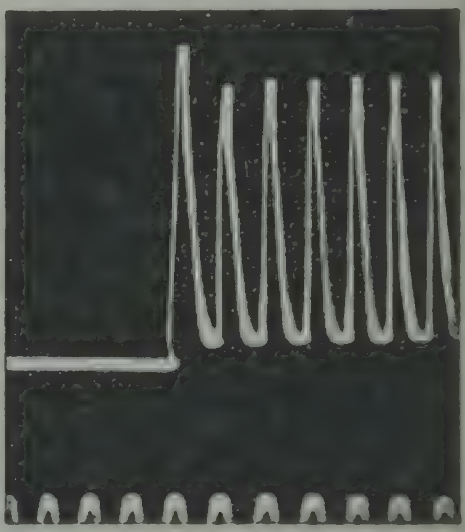


FIG. 94. Stimulation of Phrenic Nerve at 1,075/sec. All but first response are reduced. (GASSER. *J. Appl. Physics.*, 1938, **9**, 88.)

Another type of summation, which may be called *temporal summation*, exists when more than one stimulus is given to a nerve fibre. In this case, within limits to be described later, each stimulus causes a fresh impulse in each fibre, so forming *trains*, consisting of a number of impulses, equal to the number of stimuli, speeding one after the other along the fibre. This does *not* correspond with wave summation in muscle because the nerve impulses never fuse together, but always remain discrete.

By repetitive stimulation of a nerve both types would be combined, *i.e.*, there would be trains of volleys.

Alterations in Excitability. In order to stimulate a nerve, a current must have a certain minimal voltage, and, as we learnt in Chapter X, be applied for not less than a certain minimal time. These minimal values are called thresholds : the more excitable a nerve is, the smaller the threshold intensity and time needed for excitation. In order to measure excitability it is usual to make the duration of the exciting current greater than threshold time and to find the smallest voltage that will excite. It is customary to use condenser discharges for this purpose. Any change in excitability is measured by the change of threshold, and is proportional to its reciprocal. In giving results the initial excitability is called 100, and subsequent ones in percentages.

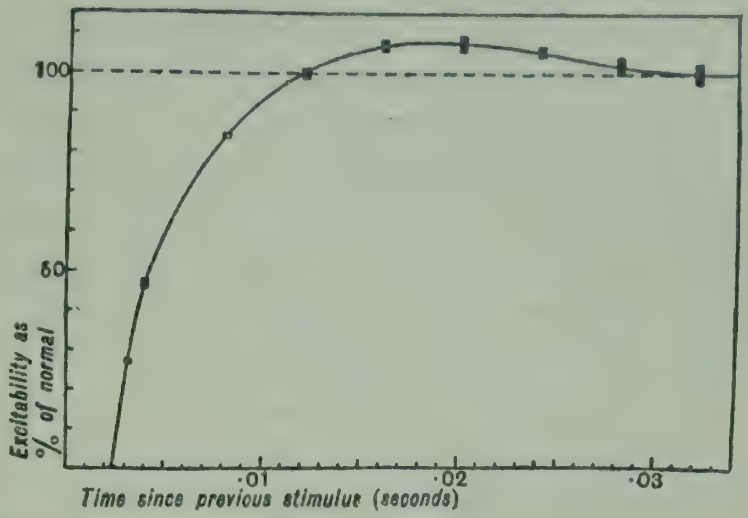


Fig. 95. Recovery of Excitability of Frog's Nerve. (Adrian and Lucas.)

Refractory Period. For a short time after its excitation a nerve fibre is in a state of altered excitability. If two "adequate" stimuli are applied to a frog's sciatic nerve within a sufficiently brief interval, the second stimulus is ineffective. This is the *absolute refractory period.* It is longer the lower the temperature and for frog nerve is about 2 msec. (0·5 msec. to 3·0 msec.) at 10° C.,* the Q10 being 3·0 (*i.e.* the duration being reduced to one-third by a rise of 10° C.). It is lengthened if the excitability of the nerve is depressed by narcotics. Following the absolute refractory period, the nerve remains in a less excitable state than normal for a further short period, during which it may only be excited by stronger stimuli than were needed for the first, or *conditioning*, stimulus, and during which time also the impulses generated are of smaller size. This is the *relative refractory period.*

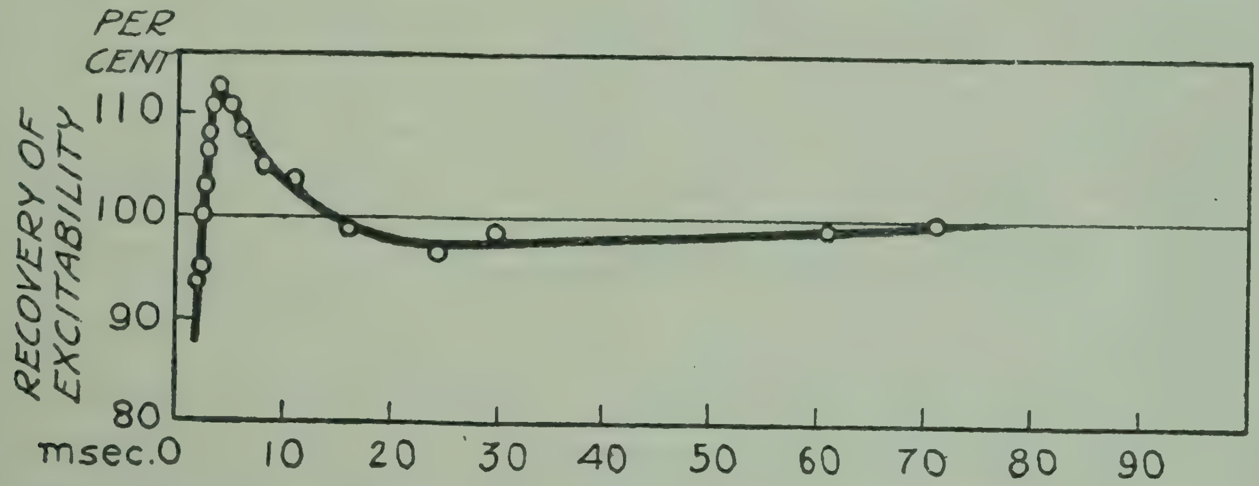


Fig. 96. Recovery of Excitability of Cat's Saphenous Nerve stimulated at 0. (Gasser and Grundfest. *Amer. J. Physiol.*, 1936, **117**, 129.)

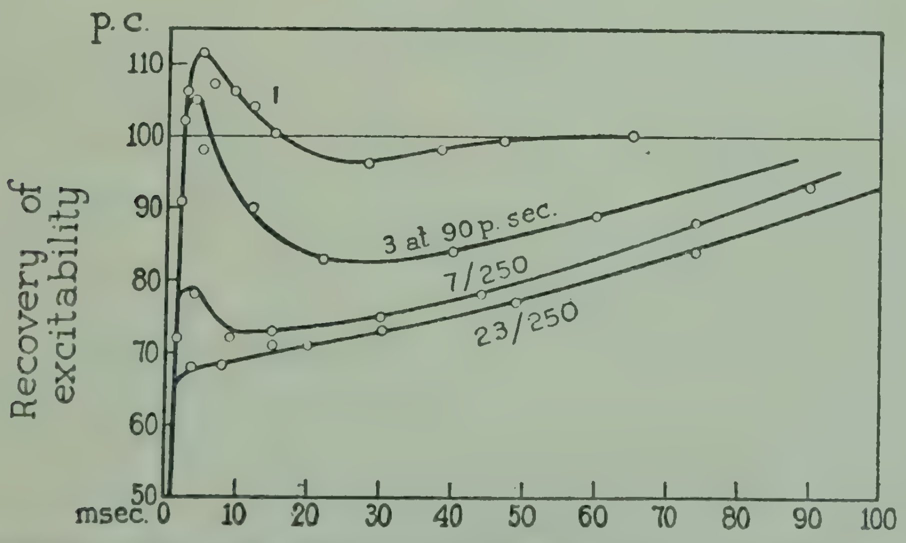


Fig. 97. Excitability of saphenous nerve of cat *in situ.* Ordinates: excitability compared with resting threshold. Abscissæ: intervals between last of the conditioning shocks and testing shock. Numbers over the curves show the number and frequency of the conditioning shocks, *e.g.* 7/250 means 7 shocks at the rate of 250 per second. The upper curve is after one conditioning shock only. (Gasser and Grundfest. *Amer. J. Physiol.*, 1936, 117.)

---

* Graham, Helen T. *Amer. J. Physiol.*, 1935, **111**, 452.

These periods of diminished excitability immediately after the application of a stimulus are succeeded by one during which the nerve is more excitable,[*] the *supernormal phase* (Fig. 96), and this again by another phase of *subnormal excitability*. For the large (A) fibres of mammals the absolute refractory period is about 0·4 msec., relative refractory phase 3 msec., supernormal phase 12 msec., and subnormal phase up to 70 msec. from stimulation (Fig. 96).

When the nerve has been submitted to several "conditioning" stimuli previous to the testing shock, the picture is somewhat altered, inasmuch as the supernormal phase and the negative after-potential are reduced or missing, the extent of the reduction being in proportion to the frequency and number of the preceding stimuli, as is shown in Fig. 97. The positive after-potential, and subnormal phase are proportionately increased.

It is an important consequence of the refractory period phenomenon that if two impulses travelling in opposite directions along the same nerve fibre meet one another, each will enter the refractory phase left behind the other, so that both will be extinguished. Another consequence is that a nerve can only carry a limited number of impulses per second. The refractory period is lengthened during continuous transmission at high frequencies.[†]

Latent Addition. The summation of two *inadequate* stimuli of, say, 95 per cent. threshold strength is possible in a nerve, provided these are not more than 0·5 msec. apart, *i.e.* the first effect is to make the nerve in the stimulated region more excitable. This is called the local excitatory state. For several msec. afterwards, however, the nerve is in a relatively refractory state, even though the stimulus was ineffective.[‡] Hence stimulation by repetition of sub-threshold stimuli may never occur if each stimulus falls in the refractory phase of its predecessor ; or if the stimuli are adequate in strength, but so rapid that (*e.g.* frog's nerve stimulated at 1,000 p. sec.), only the first one is effective (Wedensky inhibition).

Non-conducted Responses. Since the propagation of the impulse depends on the stimulation of each section of the fibre by the preceding one,[§] it should be possible to apply a localised stimulus of a strength so small that the length of nerve depolarised by it is insufficient to pass on the excitation. Hence we shall expect to get a localised and non-conducted response, and this is in fact what happens.[||] At the spot so stimulated, a second stimulus is facilitated by latent addition, and, moreover, the transitory excitation is accompanied by a small and non-propagated action potential.

All-or-None Character of the Impulse. The nervous impulse must be some form of energy transmission, since it is capable of throwing its effector organ into activity. A very important question is that concerning the type of propagation. Is it transmitted passively along the conducting medium like sound, or does the nerve fibre contribute energy which keeps the impulse alive and transmits it from point to point, as, for example, in the travelling of combustion along a train of gunpowder ?

In the former case the impulse must get progressively weaker as it travels along, and its damping-down could, under certain conditions, be greatly accelerated, as, for example, that of sound in encountering a layer of cotton wool ; in other words, the nervous impulse would normally suffer a *decrement* from which, in its further passage along its course, it would not recover.

If the latter case were the correct one, the impulse would maintain a constant (all-or-none) intensity so long as the conditions were constant and, although its progress might be damped down locally, as in a stretch of burning fuse which had been moistened, it would, if it succeeded in emerging from such a region of depression, again flare up to its normal level.

---

[*] ADRIAN and KEITH LUCAS. *Journ. Physiol.*, 1912, **44**, 68.

[†] BUGNARD and HILL. *J. Physiol.*, 1935, **83**, 394.

[‡] ERLANGER and BLAIR. *Amer. J. Physiol.*, 1931, **99**, 108

[§] HODGKIN. *J. Physiol.*, 1937, **90**, 183, 211.

[||] RUSHTON. *Proc. Roy. Soc.* (B), 1937, **124**, 210. KATZ. *Proc. Roy. Soc.* (B), 1937, **124**, 244.

Adrian * came to the conclusion that, after having passed through a region of impaired conduction brought about by local narcosis, the nerve impulse did again spring up to its former level on emerging into a normal stretch of nerve, and subsequent work has thoroughly substantiated this conclusion. That the nerve impulse normally suffers no decrement is clearly shown by the fact that the action potential as it passes along a normal nerve fibre remains of the same size. If a part of the nerve is narcotised, the action potentials are reduced in magnitude in that narcotised stretch of nerve, but if not completely quenched become normal again in the non-narcotised portions beyond. †

Adrian ‡ has pointed out that the all-or-none response of nerve is a logical consequence of its having a threshold of excitability and a refractory period. A stimulus in order to excite must produce a certain minimal local change, and the propagated disturbance which is set up as soon as this change has reached its critical value, must therefore be of an all-or-none character. A stimulus beyond threshold strength " may cause the local change to exceed the value required for excitation, but the disturbance will have been set up at the moment when the critical value was reached, and the refractory state accompanying the disturbance will prevent any subsequent change in the local conditions from affecting its size in any way."

THE EFFECT OF TEMPERATURE ON EXCITABILITY. The excitability of a nerve to currents of short duration, such as those of the spike potentials, is lowered by cooling the nerve,§ and, within certain limits, increased by raising its temperature. But to longer lasting currents the excitability is raised by slight cooling. Thus if a frog be cooled to 2° C. for a day, section of the sciatic nerve may send the gastrocnemius into continued contraction, and " closing tetanus " may be obtained with ease. This increase of " excitability " does not, however, imply a diminished chronaxie, but only a reduced rheobase. Actually the chronaxie of both nerve and muscle are increased by cooling, and therefore they are *less* excitable for induction shocks and for galvanic currents of short duration.

When a nerve trunk is warmed to a certain temperature (32° to 40° C. for various species of frogs) it becomes temporarily inexcitable (heat paralysis), but the excitability is restored again by cooling.||

The excitability and conductivity of nerve are both abolished below a certain temperature, *e.g.* 5° C. in the mammal.

THE EFFECT OF INJURY The excitability of the nerve of a nerve-muscle preparation is not equal in all parts of its course, but is greater at the upper end, probably because it is nearer to a site of injury.

Some time after a motor nerve is divided, the increased excitability at the upper end gives way to a decreased excitability, and this decrease goes on till the nerve is no longer excitable. The diminution in excitability gradually extends down the nerve fibre, so that the part of the nerve nearest the muscle remains excitable the longest (the Ritter-Valli law). It is soon followed by definite histological changes in the nerve.

PROPAGATION ALONG NERVE FIBRES

Although nerve fibres have the power of conduction in both directions *the impulse is normally only conducted in one direction*, starting at one end of the fibre and travelling to the other end. In consequence we find in the body a clear differentiation of function between various nerve fibres. Thus Bell and Magendie showed that the ventral spinal roots carry only impulses from spinal cord to periphery, while the dorsal roots carry impulses from periphery to central nervous system. The law known by the name of these observers states, indeed, that a nerve fibre cannot be both motor and sensory. We often find both kinds of fibres contained in a single nerve trunk, but normally each fibre conducts impulses only in one or other direction. (The " axon reflex," referred to later, is an exception to this, which need not be considered here.) Under normal conditions the afferent fibres are excited only at their specialised peripheral endings, while the efferent fibres are excited only at their origin within the

---

* Adrian. *Journ. Physiol.*, 1912, **45**, 413.

† Davis *et al.* *Amer. Journ. Physiol.*, 1926, **76**, 448.

‡ Adrian. *Loc. cit.*

§ Schoepfle and Erlanger. *Amer. J. Physiol.*, 1942, **134**, 694.

|| Thörner. *Z. f. allg. Physiol.*, 1920, **18**, 226.

nervous system. The difference in the function of different nerve fibres depends therefore not on the structure of the nerve fibre itself, nor on differences in the impulses conveyed, but on the different connections of the fibres. We can show this experimentally by grafting one set of nerve fibres on to another.* In the same way the finer functional differences between the various forms of sensory nerves seem to be determined by their connections within the central nervous system. Stimulation of the optic nerve by any means whatsoever evokes a sensation of light. One and the same stimulus applied to different nerves will evoke different sensations, *e.g.* a tuning-fork applied to the skin will give a sensation of vibration, to the ear a sensation of sound.

The velocity of propagation along a nerve was first measured by Helmholtz.† In a simple modification of his method, a frog's gastrocnemius is prepared, with a long piece of sciatic nerve attached. The muscle is arranged (Fig. 98) so that its contraction may be recorded on a rapidly moving surface, on which are also shown the moment at which the stimulus is sent into the nerve, and a time tracing. Records are now taken of the contraction of the muscle when the nerve is stimulated: first, at its extreme upper end; secondly, as close as possible to the muscle. It will be found that the latent period, between the point at which the stimulus is sent into the nerve and the point at which the lever begins to rise, is longer in the first case than in the second. The difference in the two latent periods gives the time that the nervous impulse has taken to travel down the length of nerve between the two stimulated points. In this way, the velocity of propagation in frog's nerve is found to be about 30 metres per second. In man and in warm-blooded animals the velocity is about 120 metres per second.‡

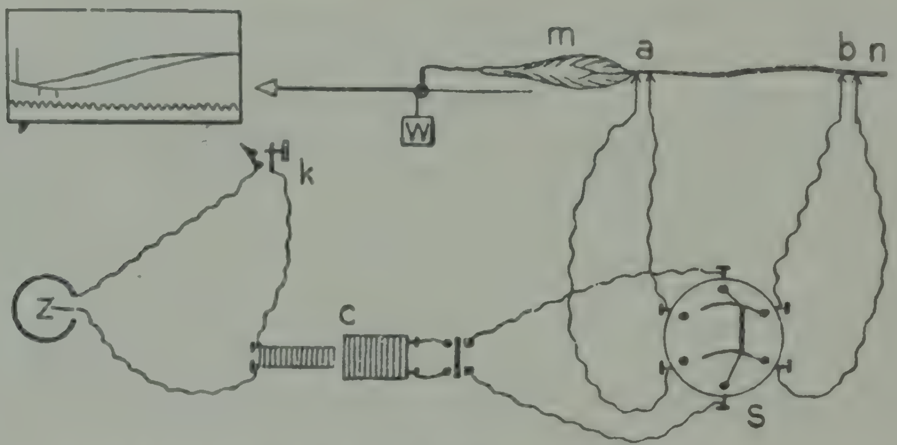


FIG. 98. Diagram of Arrangement of Experiment for the Determination of the Velocity of Transmission of a Motor Impulse down a Nerve.

The battery current passes through the primary coil of the inductorium *c*, and a "kick over" key *k*. By means of the switch *s*, the break shock in the secondary circuit can be sent through the nerve *n*, either at *b* or at *a*. The muscle *m* is arranged to write on the blackened surface, and is excited during the passage of the recording surface by the automatic opening of the key *k*. (The time marker is not shown.)

These values represent the *highest average* velocities in motor nerves. As a matter of fact, the impulses generated in a nerve of mixed function travel at different speeds, the coarse fibres conducting most quickly.§

There is a more refined method of determining the velocity of conduction, which may be used also with sensory fibres. This is by recording the times at which the monophasic potential spike passes two points at a known distance apart along the nerve.‖ We have seen that if a nerve is stimulated a volley of impulses is started simultaneously in many fibres. But, like runners starting from scratch, the farther they go the more the fast ones will outstrip the others.

---

* LANGLEY and ANDERSON. *J. Physiol.*, 1904, **31**, 365.

† HELMHOLTZ. *Arch. (Anat. u), Physiol.*, 1850, 71; 1852, 199.

‡ PIPER. *Pflüger's Arch.*, 1908, **124**, 591.

§ ERLANGER. *Harvey Lecture Series*, **22**, 1926–27, 90.

‖ SCHMITT. *J. Physiol.*, 1938, **94**, 21P.

By taking records at various distances Erlanger and Gasser found that the potentials, at first all so close together as to appear homogeneous, separated out on farther travelling into groups of different speeds, the slower groups appearing as new waves on the descending limb of the composite potential record, as illustrated in Figs. 99 and 100.

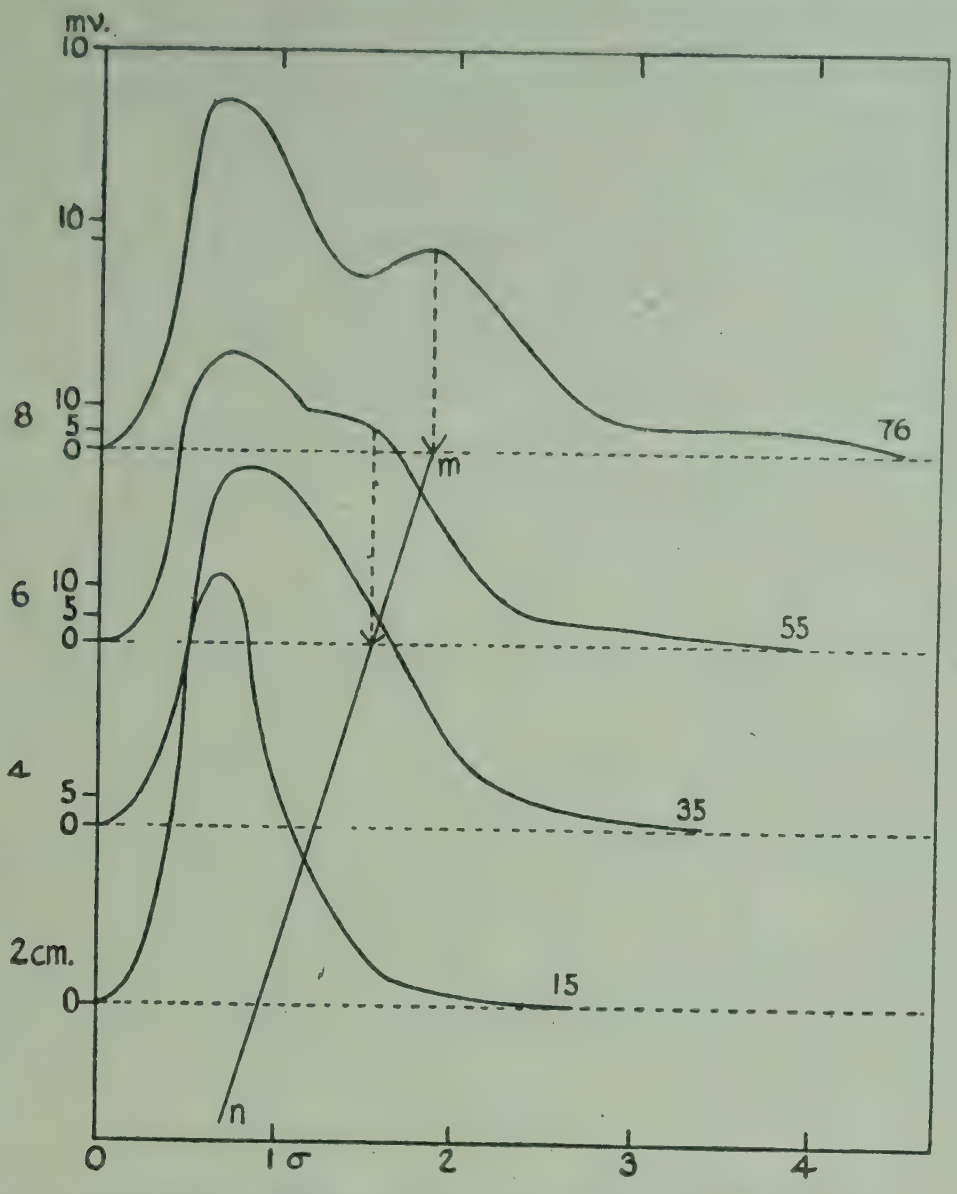


FIG. 99. Separation of Nerve Action Potentials of Frog Nerve into Components of different speeds by leading off at different distances from their origin. The figures on each curve show distance. (GASSER, ERLANGER and BISHOP.)

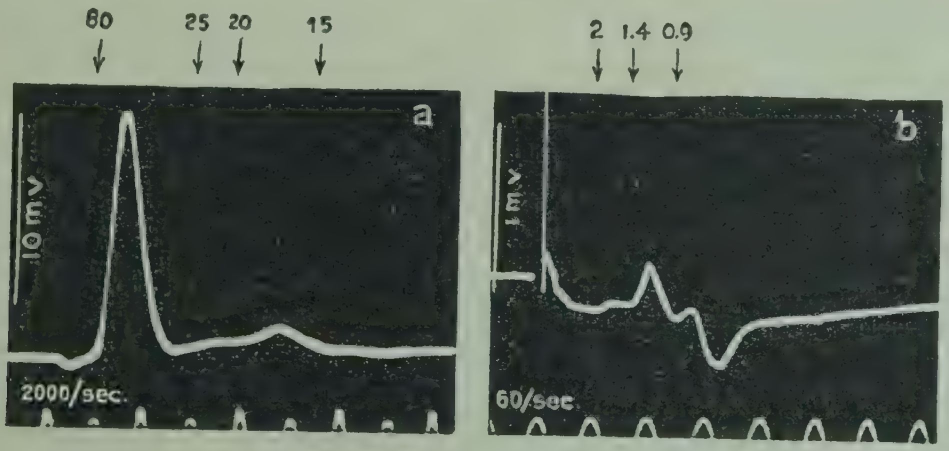


FIG. 100. Saphenous Nerve of Cat. Separation of action potentials after 4 cm. of conduction. On the left the record is on a faster "surface" than on the right and is less amplified. The spike of the left-hand figure is only faintly seen as a line in the right-hand one. The slowly recorded right-hand figure shows that the slow spikes start from a lowered line—the positive after-potential. The figures above indicate position of impulses which travel at the speeds shown (mm. p. sec.). (GASSER. *J. Appl. Physics.*, 1938, **9**, 88.)

Erlanger and Gasser have divided the mammalian fibres as follows :—

Group A are somatic medullated fibres, motor and sensory. The motor ibres are large, up to 20μ in diameter, and, together with the largest sensory bres, are called α-fibres ; these have the shortest chronaxie and the most apid propagation rate—up to 120 m.p.sec. The smaller sensory fibres (down to 1μ), β, γ and δ, have longer chronaxies, and conduction rates down to 60, 40 and 25 m.p.sec. or less ; their functions are separate.

Group B consists of autonomic medullated fibres, of fine calibre, longer chronaxie, and with conduction rates from 14 to 3 m.p.sec.

Group C consists of fine non-medullated fibres. These only conduct at 0·3 to 1·6 m.p.sec.

INTERACTION BETWEEN ADJACENT NERVE FIBRES. We have seen in Chapter X that a nerve may easily be stimulated by the action potential of a muscle placed in contact with it. This being so, the question may well be asked whether the fibres in a nerve trunk would not similarly be affected when adjacent fibres were conducting impulses. The answer seems to be * that they are affected, inasmuch as they show excitability changes, but that the effect is not enough to produce actual excitation. The experiment is carried out, with large non-medullated fibres, as shown in Fig. 101 (*a*). Fig. 101 (*b*) shows

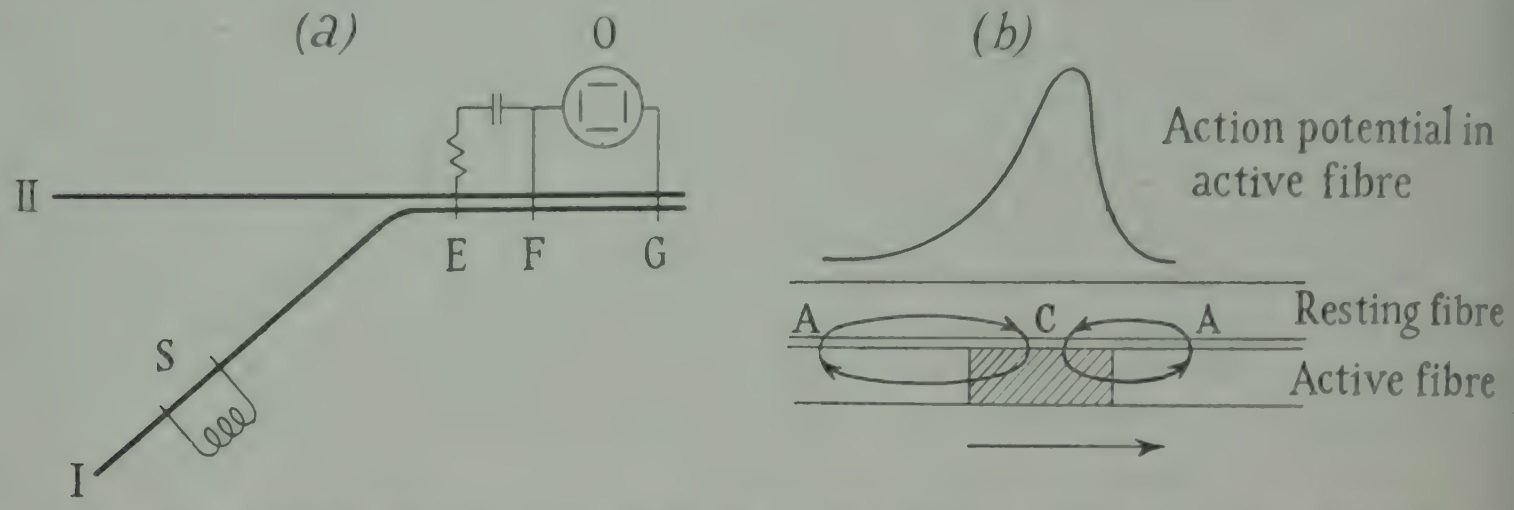


FIG. 101. Effect of action potential on adjacent resting nerve fibre. (*a*) Arrangement of experiment. I, the stimulated fibre ; II, the resting fibre. S, stimulating electrodes ; E, F, electrodes for testing threshold ; F, G, lead-off to cathode ray oscillograph, O.

(*b*) Showing that advancing action potential in active fibre results in consecutive points on resting fibre becoming successively anode, cathode, anode. (KATZ and SCHMITT.)

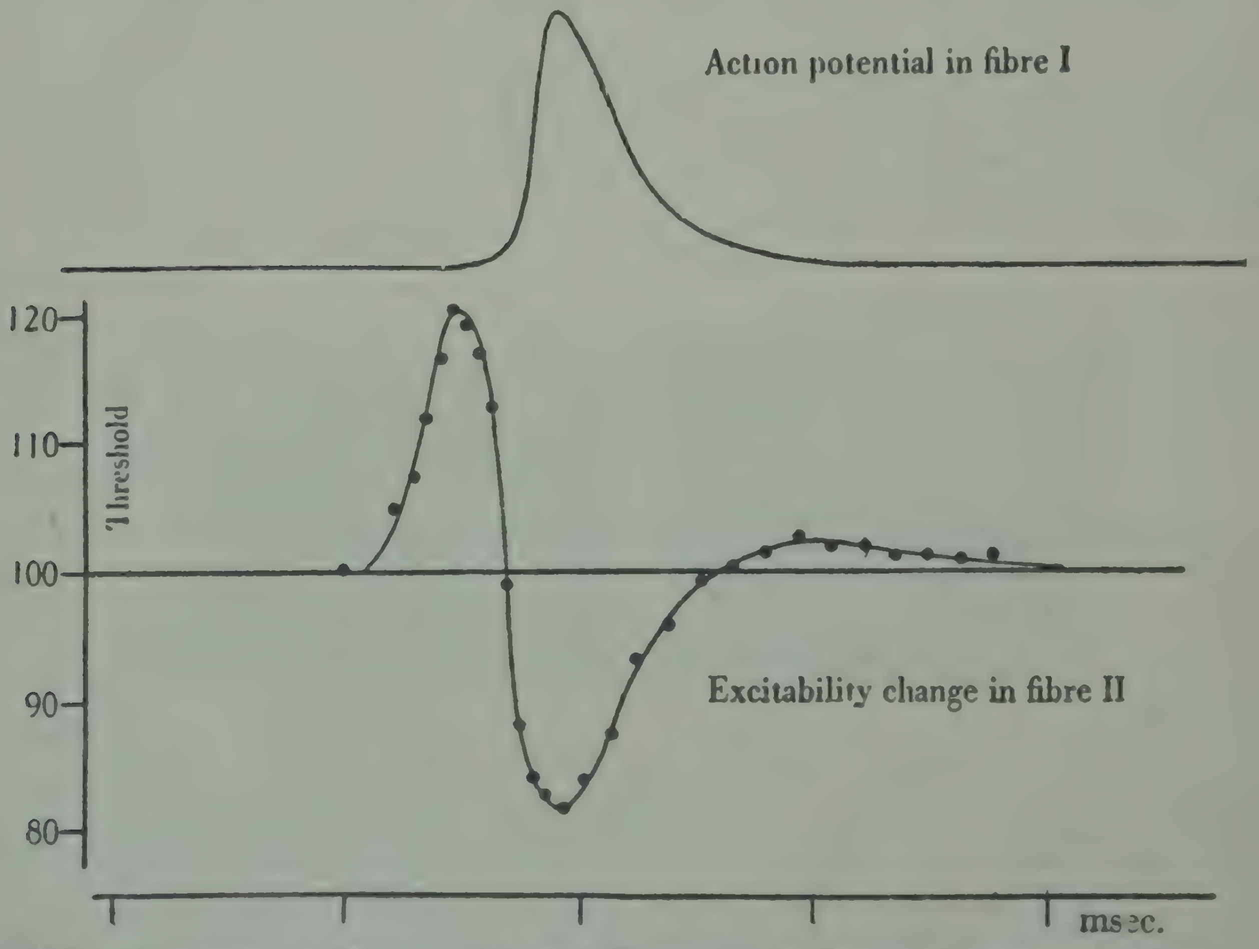


FIG. 102. Changes of Excitability of Fibre II as action potential passes along fibre I. (KATZ and SCHMITT. *J. Physiol.*, 1940, **97**, 471.)

---

* KATZ and SCHMITT. *J. Physiol.*, 1940, **97**, 471 ; 1942, **100**, 369.

that, as the action potential passes along the active fibre, consecutive areas of the resting fibre become first anelectrotonic, then catelectrotonic, and once more anelectrotonic, owing to current conducted from the active fibre. Consequently the excitability of the resting fibre is first reduced (*i.e.* its threshold is raised), then increased, and finally reduced again, as shown in Fig. 102.

There is also an effect when impulses travel simultaneously along fibres lying side by side, whereby the conduction rate of a lagging impulse is speeded up until it catches up with the other, the two thereafter keeping pace. These facts apply to non-medullated nerves, and are less conspicuous in medullated ones, because the axons are insulated.

Chemical and Thermal Changes in the Nerve during Transmission. In nerve, which serves merely as a conducting medium, there is less expenditure of energy on excitation than in muscle. All that is necessary is that each section of the nerve should transmit to the next section just so much energy as is needed to excite it.

That the excitatory process in nerve is dependent upon chemical changes is indicated by the facts that, in the complete absence of oxygen, the nerve fibres lose their excitability, and that this loss of excitability is hastened by repeated stimulation of the nerve. When the excitability has been abolished by stimulation in the absence of oxygen, it may be restored within a few minutes by re-admission of oxygen to the nerve, and still more quickly if the nerve be also washed in Ringer's solution.

Nerve continually uses oxygen and evolves carbon dioxide ; in the absence of oxygen, lactic acid accumulates.* When the nerve is stimulated, the rate of oxygen usage is accelerated.† The substances oxidised are not known. Little is known concerning actual chemical changes in nerve. Carbohydrate is apparently not the main source of energy.

The metabolism of nerve cells is apparently entirely different, for the brain appears to use carbohydrate exclusively ; if glucose is absent from the blood, the brain functions are grossly disturbed.

In activity acetyl choline is released by cholinergic fibres and sympathin by adrenergic ones. Potassium ions leave the fibre and sodium ions enter. The vitamin thiamine has also been shown to leave the nerve during activity.‡ Breakdown of A.T.P. and creatine phosphate also probably occur.

Observations of heat production of nerve by A. V. Hill and his pupils also show how small are the energy exchanges. Frog's nerve at rest evolves about 4×10^{-3} cal./g./min. when in oxygen, and about one-quarter of that amount when in nitrogen.§ If returned to oxygen after being in nitrogen, the heat production is increased.

When impulses pass along the nerve, the heat production is increased, and three phases may be recognised,‖ the first being called "initial" heat and the other two "recovery" heat. Initial heat is small in amount, *e.g.* 5 to 10 micro-calories (*i.e.* $5–10 \times 10^{-6}$ cal.) per sec. per g. nerve during continuous rapid excitation. It dies away at once when stimulation ceases. The delayed heat of recovery consists of an early rapid phase, about equal to the initial heat, and a prolonged phase for 30 minutes or more, which totals about thirty times the initial heat. The heat produced by a single volley of impulses is of course much smaller, viz. about 0·067 micro-calories per g. of frog nerve at 20° C.

An explanation of these various facts is as follows : ¶ The nervous impulse is a wave of altered polarisation which rapidly moves over the polarised membranes at the surfaces of the fibres, thus also causing the action potential. The altered polarisation means that ions previously held in position on the polarised membrane are now free to wander. This change would involve but little liberation of energy and so necessitate little if any heat production. Hence it occurs equally well at first whether oxygen is present or not. But the process of *recovery* involves *repolarising* the membranes, which means that energy must be expended, oxygen used and heat set free, in order to separate the oppositely charged ions and set them in order again. This explains why these various changes are associated with the recovery of the nerve rather than with its actual functioning, and also why the excitability is lost when the nerve is stimulated repeatedly in absence of oxygen, and restored by its re-admission. At the *anode* of a constant current, the effect is the same as that of oxidative metabolism, *i.e.* the polarised state is restored.

The hypothesis also explains another phenomenon called the "retention of action current " which is well seen in cold blooded non-medullated nerves.[*] When the nerve is repeatedly stimulated, especially in absence of oxygen, the action current subsides more and more slowly, or, in other words, the depolarised state persists, as a negative after-potential. If oxygen is admitted the current subsides more quickly again, because the energy necessary to repolarise the membranes is obtainable. The more rapidly acting medullated nerves of the higher animals do not show the phenomenon so clearly, but it is demonstrable in them too. The refractory period is, indeed, due to the fact that repolarisation takes an appreciable time.

---

* GERARD and MEYERHOF. *Bioch. Ztschr.*, 1927, **191**, 125.
† GERARD. *Amer. J. Physiol.*, 1927, **82**, 381 ; FENN. *J. Gen. Phys.*, 1927, **10**, 767.
‡ v. MURALT. Die Signalübermittlung im Nerven. 1946. (Birkhaüsen, Basel.)
§ BERESINA. *J. Physiol.*, 1932, **76**, 170.
‖ HILL. *Proc. Roy. Soc.*, 1933, **113B**, 345.
¶ HILL. *Proc. Roy. Soc.*, 1929, **105B**, 153.

For Reference

Erlanger (1939). The Initiation of Impulses in Axons. *J. Neurophysiol.*, **2**, 370.
Gasser (1939). Axons as Samples of Nervous Tissue. *J. Neurophysiol.*, **2**, 361.
Gerard (1932). Nerve Metabolism. *Physiol. Rev.*, **12**, 469.
Feng (1936). The Heat Production of Nerve. *Ergebn. d. Physiol.*, **38**, 73.
Page (1937). Chemistry of the Brain. (Baillière, Tindall & Cox, London.)
Katz (1939). Electric Excitation of Nerve. (Oxford Univ. Press.)

CONDITIONS AFFECTING THE CONDUCTIVITY OF NERVE

Temperature. Below a certain temperature the propagation of the excitatory process in the nerve is abolished. In the frog it is necessary to cool the nerve to − 1° C. before conduction is abolished, whereas in the mammal it is sufficient to cool the nerve to somewhere between 0° and 5° C. Since cooling the nerve does not excite it, this procedure forms a convenient method for blocking the passage of impulses along a nerve without using the irritating procedure of section. On again warming the nerve the conductivity returns. Warmth (over 32° C.) abolishes conductivity in frog nerve. The rapidity with which the excitatory process is propagated along nerve fibre depends on the temperature. Thus, the mean rate of conduction in the frog's nerve at 8° to 9° C. is only 16 metres per second. The temperature coefficient of the velocity of nerve propagation, *i.e.* $\frac{\text{velocity at T} + 10^\circ}{\text{velocity at T}^\circ}$, is about 1·58 between 12° and 26°, but below 12° is higher.[†]

Pressure. Local application of pressure above 130 mm. Hg to a nerve produces in about 40 min. a block that lasts for 2–3 hours after removal of the pressure : this differs from anæmic block which disappears when the circulation is restored.[‡]

Chemical Influences. The most important drugs with an influence on nerve fibres are the narcotics. Of these we may mention carbon dioxide, ether, chloroform, alcohol and novocaine.

These diminish, and finally abolish, the excitability and conductivity of the nerve fibres. On removing the narcotic, recovery occurs if the dose has not been too great. It has been shown that the nervous impulses (action potentials) are reduced in magnitude while passing through a lightly narcotised stretch of nerve, but regain their normal size on emerging into untreated tissue.[§] If the narcosis is severe any impulses are extinguished altogether so long as it lasts. Advantage is taken of this fact for causing a temporary loss of function in nerve trunks for clinical purposes. Thus, if novocaine is applied to a nerve trunk, all painful sensations from the part supplied by the nerve are blocked, and operations may thus be performed under " local " or " regional " anæsthesia. If applied to the spinal cord (by lumbar puncture), extensive operations on the abdomen or legs can be carried out painlessly.

Changes in hydrogen-ion concentration do not affect the spikes, but alter the excitability and the after-potentials. Increase of *c*H reduces excitability and reduction of *c*H increases it, until at *p*H 8·0 spontaneous discharges occur.[‖]

Removal of calcium produces the same effects as reduction of *c*H, which is, in fact, probably due to de-ionisation of calcium. Removal of potassium produces the opposite effects.[¶]

---

[*] Furusawa. *J. Physiol.*, 1929, **67**, 325.

[†] Rosenberg and Sugimoto. *Bioch. Ztschr.*, 1925, **156**, 262.

[‡] Bentley and Schlapp. *J. Physiol.*, 1943, **102**, 72.

[§] Adrian. *J. Physiol.*, 1913, **45**, 389 ; Davis, Forbes, Brunswick and Hopkins. *Amer. J. Physiol.*, 1926, **76**, 448.

[‖] Lehmann. *Amer. J. Physiol.*, 1937, **118**, 600.

[¶] Lehmann. *Ibid.*, p. 613.

Asphyxia reduces the size of the response, *i.e.* of the spike and after-potentials, and also reduces excitability.[*]

Block Phenomena. When a short length of nerve is blocked, *e.g.* by cold or pressure, and an impulse started above the block, this is not entirely without effect on the nerve below the block, although no impulse has passed. For a few millimetres below the block the nerve shows for a few msec. a large increase of excitability, the threshold being often reduced by 80 per cent. Thus in an experiment arranged as in Fig. 103 a shock at S_1 sets up an impulse which is stopped at the block and so does not reach the oscillograph C. A sub-threshold shock from S_2 similarly fails to arouse an impulse. But if an adequate shock S_1 is followed 1 or 2 msec. later by a sub-threshold shock at S_2, an impulse is started and is recorded at C.

The explanation, which is of considerable theoretical interest, is that, although the impulse cannot be conducted beyond the block because of the alterations produced by the conditions, the electrical disturbance due to the impulse above the block causes currents which can be conducted perfectly well. The direction of these currents is shown in Fig. 104 from which will be seen that the current below the block is in the same direction as would be got if a cathode were applied to the nerve. Measurements at the nerve surface here do, in fact, show that an " extrinsic " potential is set up which is similar in extent and duration to one set up by applying a cathode, and it is known that the region of the cathode is one of enhanced excitability. The extrinsic potential is therefore electrotonic in nature, and resembles the non-propagated response already mentioned.

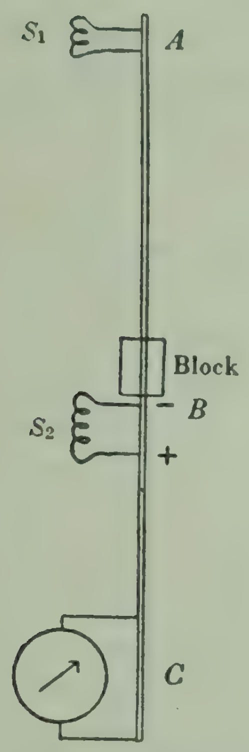


Fig. 103. Experiment to show excitability change beyond Block. (Hodgkin. *J. Physiol.*, 1937, **90**, 183.)

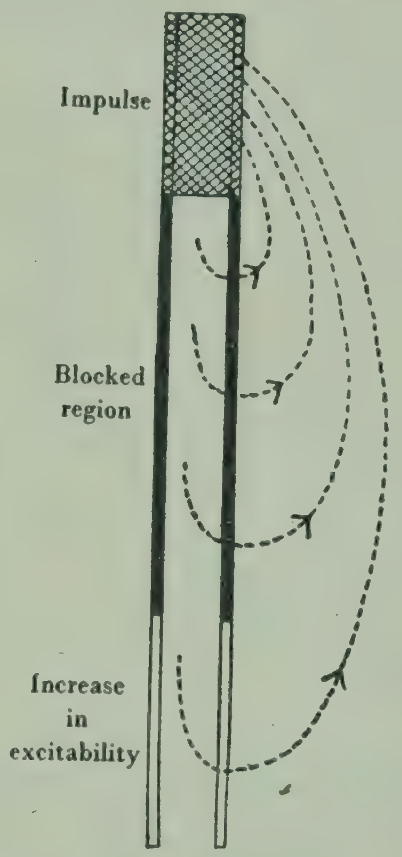


Fig. 104. Diagram of electrical circuit set up in blocked nerve by an impulse above the block. (Hodgkin. *J. Physiol.*, 1937, **90**, 184.)

The Influence of Repeated Stimulation. Medullated nerves are not at all easily fatigued, in fact they are almost indefatigable unless deprived of oxygen. This may be shown in various ways. Thus, a nerve can be isolated and repeatedly stimulated, say, at 100 per second, while its action potentials are observed on a cathode ray tube ; they will be found to continue unchanged for long periods. It may be demonstrated in mammals by poisoning the animal with curare, and then stimulating a motor nerve continuously while the animal is kept alive by means of artificial respiration. As the effect of the curare begins to wear off in consequence of its excretion, the muscles supplied by the stimulated nerve enter into tetanus. The action of the curare may be cut short at any time by the injection of eserine, when the muscles will at once begin to react to the excitation.

---

[*] Lehmann. *Ibid.*, 1937, **119**, 111.

The same fact may be shown in the frog. The gastrocnemii of the two sides, with the sciatic nerves, are dissected out, and an exciting circuit is so arranged that the interrupted secondary currents pass through the upper ends of both nerves in series (Fig. 105). A galvanic cell is connected with two non-polarisable electrodes (*np*, *np*) placed on the nerve of B, so that a current runs in the nerve in an ascending direction; this will block the passage of impulses. The exciting current is then sent through both nerves by the electrodes e_1 and e_2. The muscle A enters into tetanus, which gradually subsides owing to "fatigue."

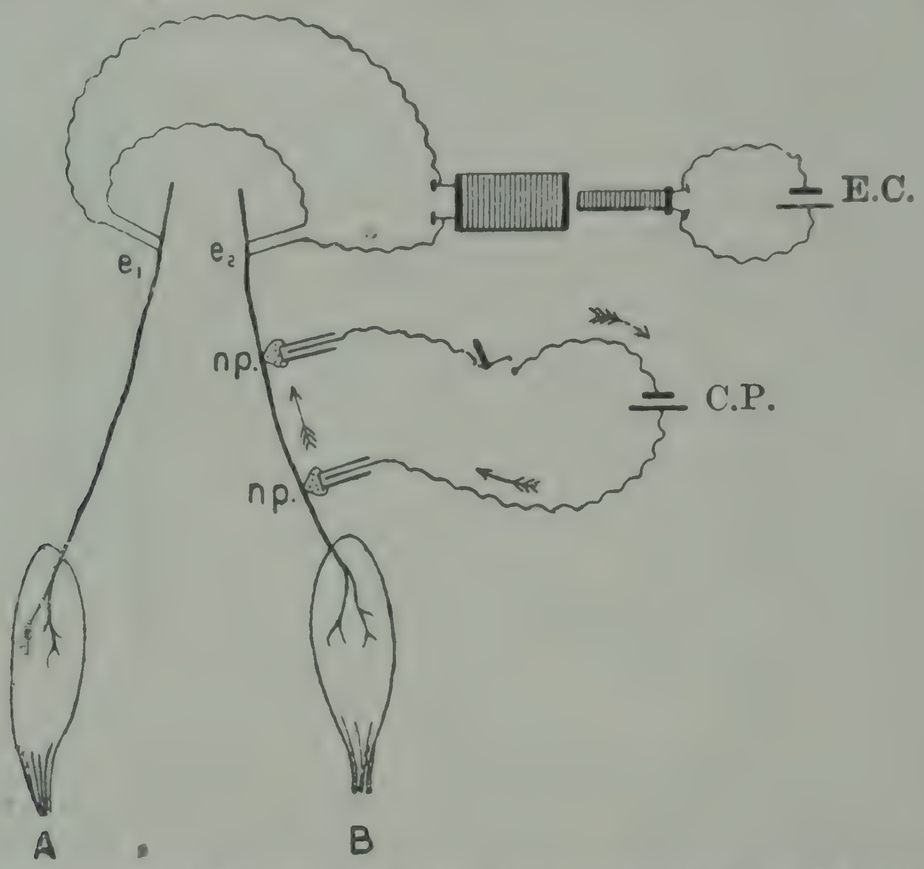


Fig. 105. Arrangement of Experiment for demonstrating the Relative Absence of Fatigue in Medullated Nerve Fibres.
E C, exciting circuit; C P, polarising circuit.

When A no longer responds to the stimulation, the constant current through the nerve of B is broken. B at once enters into tetanus. Since both nerves have been excited continuously it is evident that the fatigue does not affect the nerve trunk.

But although repeated stimulation does not reduce the nerve to an inexcitable and non-conducting state it would be wrong to suppose that it produces no change at all. The phenomena of the refractory period and the after-potentials are indications of incomplete recovery of normal membrane polarisation, and hence if stimuli follow one another with sufficient rapidity, certain consequences follow. If the intervals between stimuli are shorter than the relative refractory period, the second and subsequent stimuli produce smaller spikes than the first one (*v.* Fig. 94), and may show a progressive decline. Also as stimulation continues the negative and positive after-potentials increase (Fig. 93), and during the positive after-potential the excitability and conduction velocity are subnormal.*

ELECTROTONUS AND POLARISATION PHENOMENA

Electrotonus. When a constant current is passed along a motor nerve its direction is called *descending* when the anode is the electrode further from the muscle, and ascending when it is the nearer. Such a constant current is often called a *polarising current*, and many of the phenomena produced by it are due to changes of the state of polarisation of certain structures in the nerve. If the nerve be very excitable, the muscle may be thrown into a tetanus during the whole time that a *strong descending* current is passing ("closing tetanus"). On the other hand, if a strong *ascending* current be passed through the nerve for a considerable time, the muscle when the current is broken may go into "opening tetanus" for several seconds. Usually, however, the muscle responds with a single twitch at the make and break of a moderate current, although, on investigating the condition of the nerve *during* the passage of the current, we find that it is considerably modified. This modification in the condition of the nerve is spoken of as *electrotonus*, and includes changes in its excitability, conductivity and electrical condition.

---

* Gasser. *Amer. J. Physiol.*, 1935, **111**, 35.

To demonstrate the changes in *excitability* the apparatus shown in Fig. 106 may be used, with a nerve muscle preparation. A constant current from the battery is led through a part of the nerve by means of non-polarisable electrodes, about one inch apart. In this circuit we put a reverser and a key. This is the polarising circuit. The coil is arranged so that we may use break induction shocks, which are applied to the nerve to test its excitability, in the neighbourhood of the pole nearest to the muscle.

We employ submaximal induction shocks throughout the experiment. We now throw a descending polarising current into the nerve, *i.e.* the cathode is nearer to the muscle. At the make of this current the muscle will respond with a twitch which we may ignore. We now, with the current still flowing, test the excitability of the nerve, and we find that in the region of the cathode the stimulus, which before was submaximal, has now become maximal.

We now reverse the direction of the polarising current, and find that the stimulus when applied near the anode has little or no effect (Fig. 107).

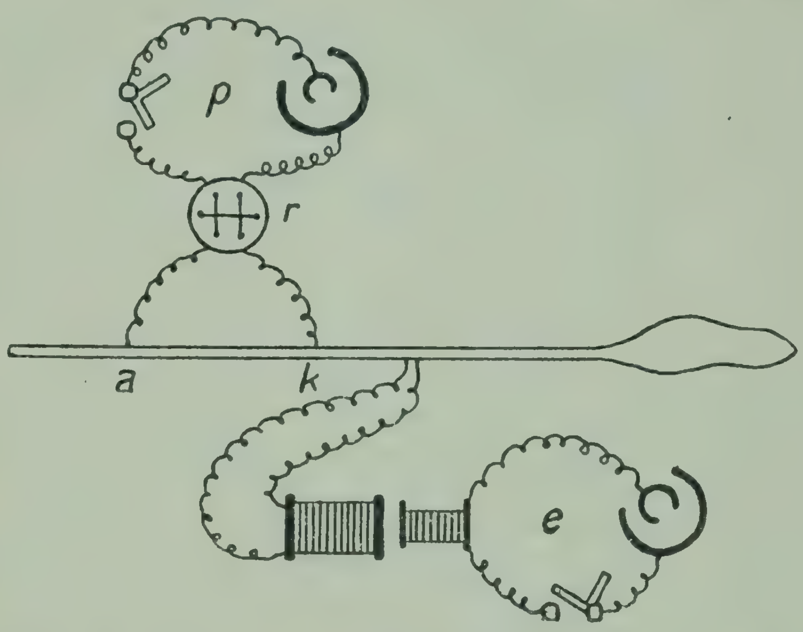


FIG. 106. Arrangement of Apparatus for showing Electrotonic Changes in Excitability. *e*, exciting current; *p*, polarising current; *r*, Pohl's reverser.

If the testing stimulus be applied between the anode and cathode, we obtain information as to the *conductivity* of the nerve as a result of electrotonus. We thus find that in anelectrotonus the impulse may be completely blocked if the polarising current is a strong one (anelectrotonic block).

On break of the polarising current the condition of the nerve returns to normal, and the stimulus is once more submaximal throughout. This return to normal is not immediate, since the first effect of breaking the current is a swing back, so to speak, past the normal, the diminished excitability at the anode giving place to an increased excitability which gradually subsides, while the neighbourhood of the cathode acquires a condition of diminished excitability, and this only gradually gives place to a normal condition.

This experiment teaches us that, during the passage of a constant current through a nerve, there is an increase in the excitability in the nerve near the cathode, and a diminution in excitability and conductivity near the anode. These conditions of increased and diminished excitability are spoken of as *catelectrotonus* and *anelectrotonus* respectively. Just after the breaking of the current there is a reversal of these states, called the *resolution of catelectrotonus* (or of anelectrotonus). Now the event that takes place at the cathode on make, and at the anode on break of a constant current is, as the last experiment shows us, a rise in excitability, in the former case from normal to above normal, in the latter from subnormal to normal. Hence we may say that the excitation by a constant current is caused by a sudden rise of excitability,* culminating in the discharge of an impulse, which may be due either to a sudden appearance of catelectrotonus or a resolution of anelectrotonus. We have said sudden because the steepness of the change is, as we have seen elsewhere, a necessary factor in causing excitation. The more excitable the tissue, *i.e.* the shorter its chronaxie, the more rapid must the change be if it is to culminate in excitation. In muscle we have seen similarly that a make contraction always starts from the cathode, and a break contraction from the anode.

---

* ROSENBERG. *Proc. Roy. Soc.* (B.), 1937, **124**, 308.

There must be, somewhere between the anode and cathode, an indifferent point—that is to say, a region where the excitability is neither increased nor diminished. This indifferent point is nearer the anode when the polarising current is weak; and approaches the cathode as the current is strengthened, so that with very strong currents nearly the whole intrapolar length is in a condition of anelectrotonus (Fig. 107).

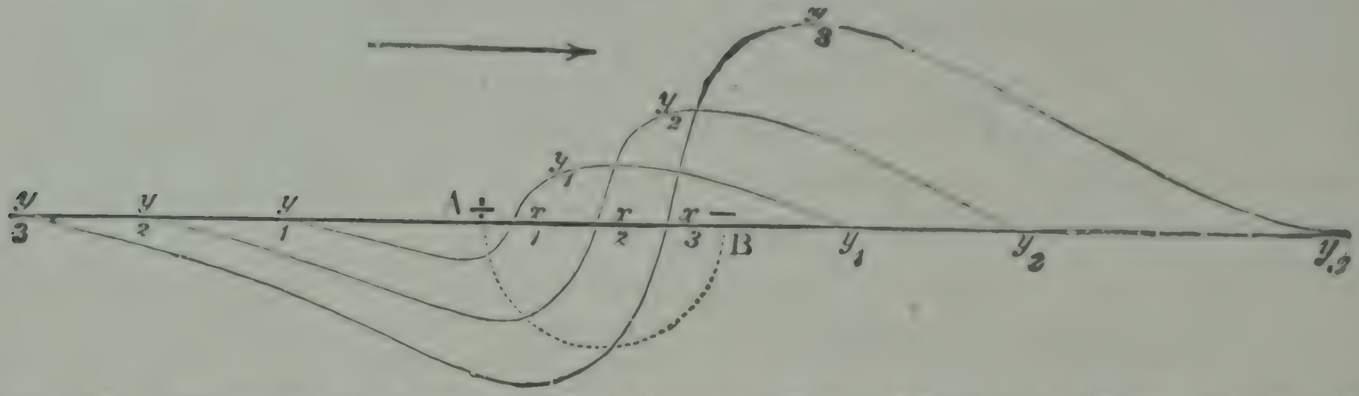
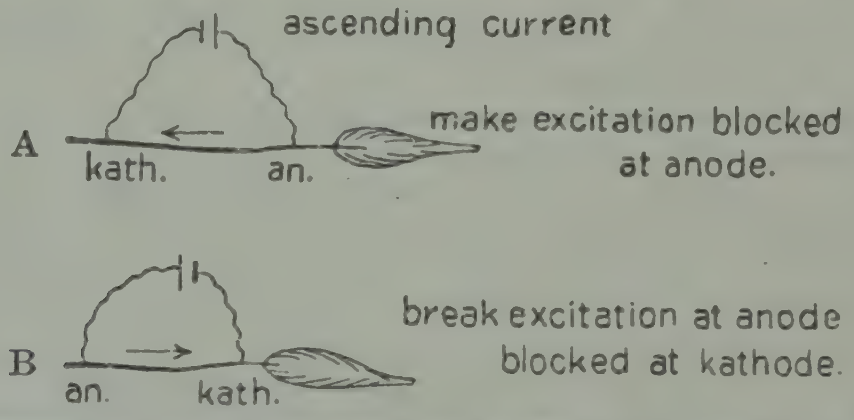


FIG. 107. Diagram to show the Variations of Excitability in a Nerve during the Passage of Polarising Currents of different Strengths. The degree of change is represented by the distance of the curves from the base line; the part of the curve below the line signifying *decrease*, that above the line *increase* of irritability.

A, anode; B, cathode; y_1, effect of weak current; y_2, medium current; y_3, strong current. It will be noticed that the indifferent point, x, where the curve crosses the horizontal line, approaches nearer and nearer the cathode as the current is increased in strength. (After PFLÜGER.)

When a strong polarising current is used, the depression of conductivity at the anode is so marked that no impulse can pass this region. Thus if we send a very strong ascending current through the nerve, there is no contraction at make. This is owing to the fact that the impulse started at the cathode on make of the current cannot reach the muscle, its passage down the nerve being blocked in the region of the anode (Fig. 108, A). Similarly at break of a strong descending current, the excitation originating at the anode is blocked by the great depression at the former cathode (Fig. 108, B).



ascending current
A
kath.
an.
make excitation blocked at anode.
B
an.
kath.
break excitation at anode blocked at kathode.

FIG. 108. Diagram to show the Blocking Effect of a strong Constant Current passed through the Nerve of a Nerve-muscle Preparation.

With the weakest currents excitation occurs only at make, since the rise of catelectrotonus, is always more effectual than a break stimulus, *i.e.* the resolution of anelectrotonus.

These results probably hold good for all excitable tissues, and, when tabulated to indicate the results of stimulating nerve-muscle preparations by opening and closing circuits with weak, medium or strong currents, ascending or descending, are sometimes spoken of as "*Pflüger's law of contraction.*"*

If the electrodes are less than 20 mm. apart, the current necessary to excite a nerve is greater the nearer they are to one another. The threshold current leaving the cathode is the same whatever the interpolar length.† But, owing to the greater electrical conductivity of the interstitial fluid as compared with the sheaths of the axis cylinders, the nearer the electrodes are together the greater the fraction of the total currents shunted by the interstitial fluid.

ELECTROTONIC POTENTIALS. If points *e* near one of the electrodes P be tapped off during the passage of a polarising current, and their potentials compared with that of

---

* PFLÜGER. Unters. ü. d. Physiol. des Electrotonus. 1859. (Berlin: *Hirschwald.*)

† RUSHTON. *J. Physiol.*, 1934, **82**, 332.

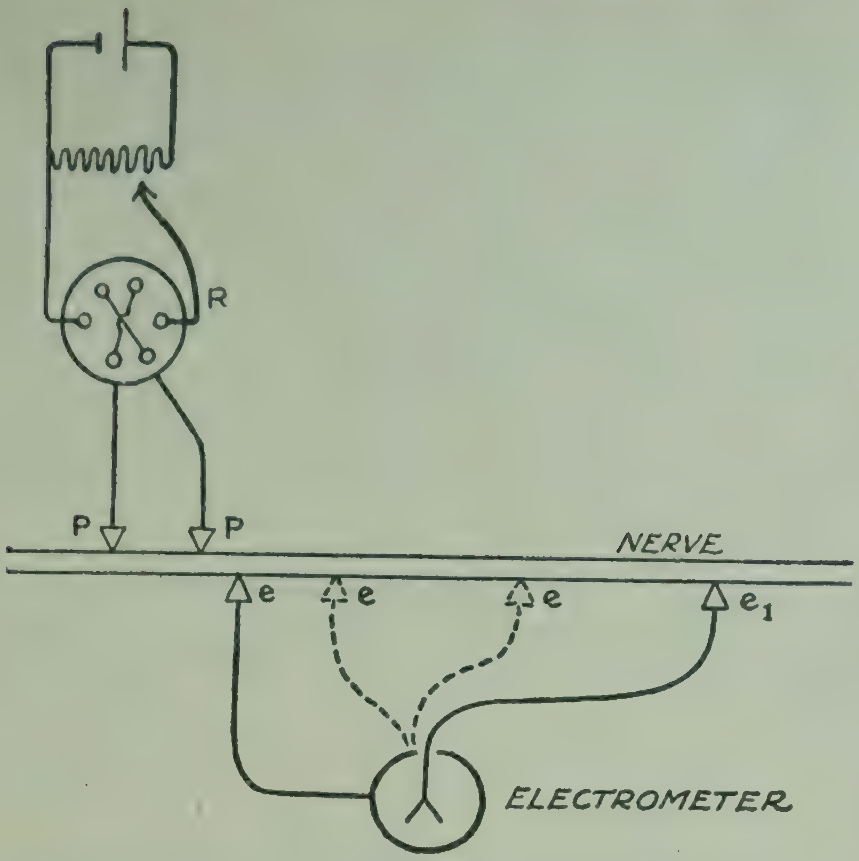


FIG. 109. Measurement of Electrotonic Potentials. (After HECHT.)

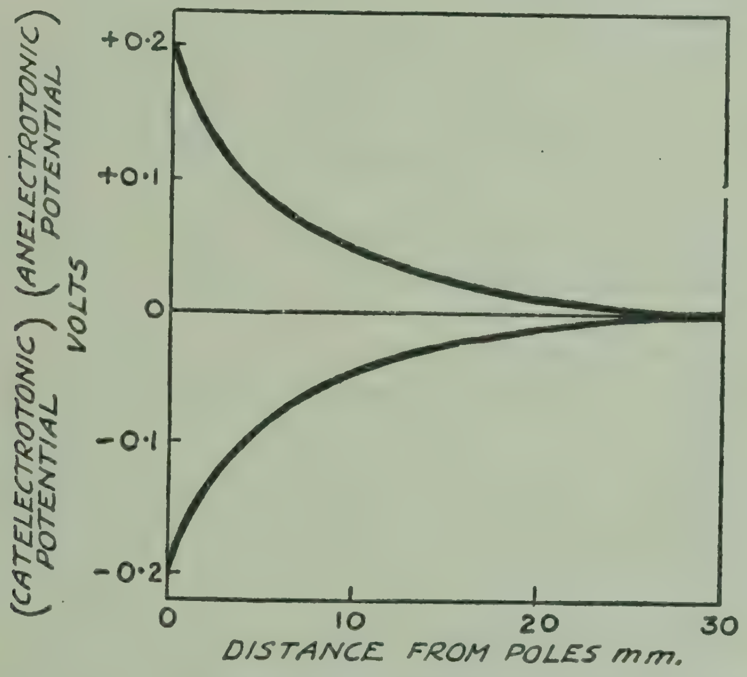


FIG. 110. Distribution of Electrotonic Potentials.

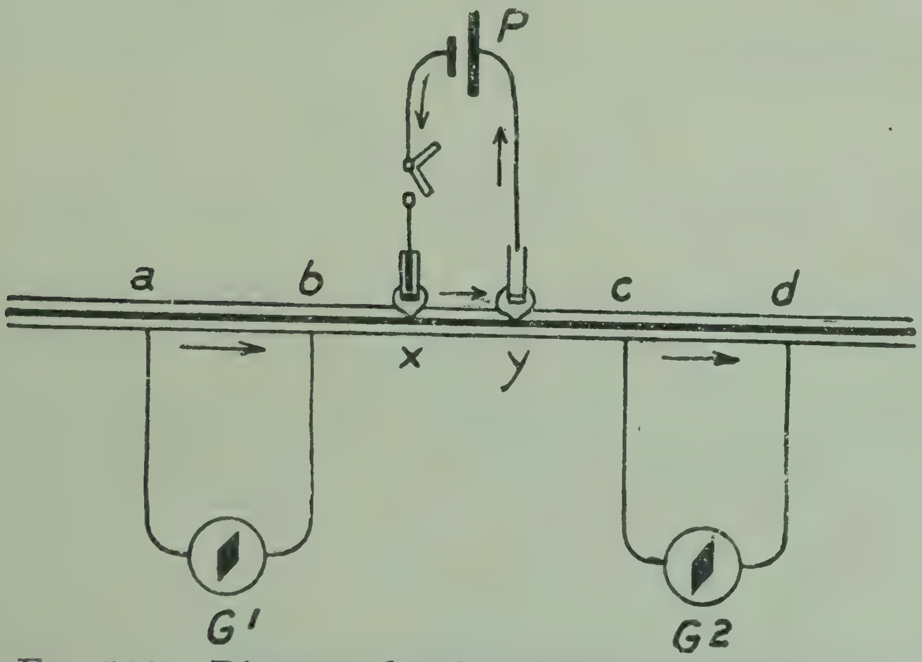


FIG. 111. Diagram showing Electrotonic Currents.
P, polarising circuit; G_1, G_2, galvanometers.

a more distant point e_1 on the nerve (Fig. 109) by an electrometer, it is seen that the local potentials extend for some distance beyond the electrodes* (Fig. 110). These are called *electrotonic potentials*, and are confined to the outer sheaths of the structures concerned.

---

* HECHT. *Z. f. Biol.*, 1931, **91**, 252.

Electrotonic Currents. If, instead of an electrometer we use a galvanometer through which a current can pass, between extrapolar points, as illustrated by *a*, *b*, *c*, *d*, in Fig. 111 we find a steady current in the extrapolar portions of the nerve from *a* to *b*, and from *c* to *d* *i.e.* *both* currents are in the *same* direction along the fibre as the polarising current.

The currents so produced are called the "electrotonic currents." They must not be confounded with the action potential which originates at one of the poles only at make or break of the current, and is transmitted thence as a wave. The electrotonic potential, on the other hand, commences almost instantaneously,* takes a very brief time to reach its maximum, then remains nearly stationary, lasts the whole time that the current is flowing through the nerve, and does not spread far along the fibre. It is caused by the occurrence of polarisation in membranes lying between the sheath and the better conducting central part of the nerve fibre, and may be reproduced on a model made up of a glass tube filled with normal salt solution, containing a wire stretched through it (Fig. 112).† The polarisation in question occurs at the surface of the wire. On leading a current through *a* and *b*, and connecting *c* and *d* or *e* and *f* with a galvanometer, a current will be observed in the extrapolar portion of the model which, as in nerve, is in the same direction as in the intrapolar. A useful model illustrating the electrical properties of nerve may also be made with resistances and leaky condensers.*

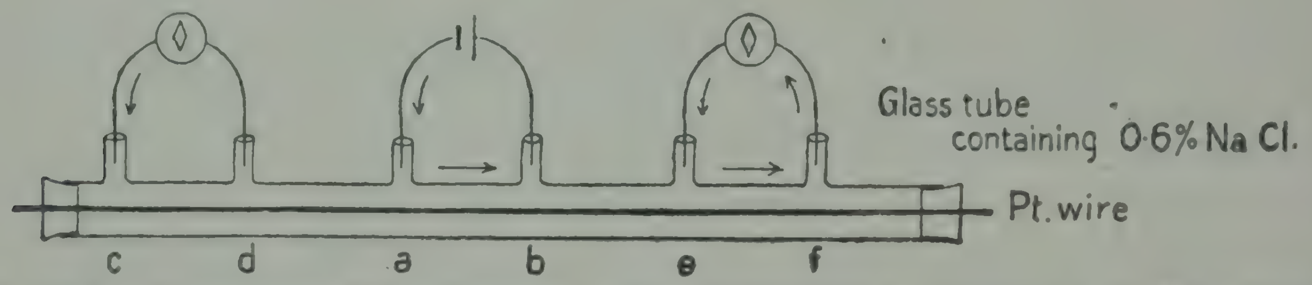


Fig. 112. Apparatus for imitating the Polarisation Phenomena in Medullated Nerve ("Core-conductor" model).

In the case of nerve, which normally has polarised membranes, the effect of the constant current is an increase of the natural polarisation beneath the anode and a decrease beneath the cathode.

This altered polarisation will not disappear at once on breaking the polarising current. On replacing the battery by a galvanometer, we shall get a current through it in the direction opposite to the previous polarising current, viz. from anode to cathode.

Relation of Electrotonic Changes to Excitation. When a constant current is led into a nerve the potential at the cathode rises very rapidly from the resting value V_0, by the establishment of an electrotonic potential, V. The rise of electrotonic cathode potential is exponential, and when it has reached a certain value, say U_0, the extent of the excitation due to polarisation change becomes sufficient to be self-propagating. With a current of rheobasic strength this is the fullest polarisation the current can produce; if of twice rheobase, excitation occurs when half the possible polarisation has occurred, and so on. In other words, for currents of short duration a constant minimal quantity of electricity is needed to produce excitation. This quantity is about 10^{-9} coulombs for medullated nerve and about 10^{-6} for non-medullated nerves.‡

The time constant *k* of the change of cathode potential from V_0 to U_0 in the expression

$$-\frac{dU_0}{dt} = \frac{(U_0 - V_0)}{k}$$

has a value for frog's nerve at 20° of about 0·35 msec., and is related to the chronaxie, which $= 0{\cdot}693\,k$.§

A second time-factor has been observed by Hill, and is called the *time-factor of accommodation.* It is well known that a slowly rising current has a higher threshold than an abruptly increasing one—the tissue seems to get accustomed to the stimulus and can disregard it until it reaches a greater strength before it responds to it. Let us call the potential at this raised threshold U_x. Then, on passing the rising current the threshold of the nerve has risen from U_0, its former threshold, to U_x at a definite rate, given by the expression:

$$-\frac{dU_x}{dt} = \frac{(U_x - U_0)}{\lambda}$$

where λ is the time-constant of accommodation. Its value is from ten to one hundred times as great as that for *k*, *i.e.* it is about 35 msec. for frog's nerve: it may vary independently of *k*, the most remarkable instance of this being produced by alteration in the concentration of Ca ions.* If the Ca ions are increased the value of λ is greatly reduced, *i.e.* the tissue accommodates itself very rapidly, or, in other words, it takes more sudden changes to excite it. Conversely, when the Ca ions are reduced, the tissue becomes hyper-excitable; this is a well-known fact in connection with the occurrence of tetany consequent on parathyroidectomy, or over-breathing, which reduce the Ca-ion concentration of the blood.

---

* Bogue and Rosenberg. *J. Physiol.*, 1934, **82**, 353.

† For an account of these classical models see Gotch in Schäfer's Text-Book of Physiology, 1900, Vol. II, p. 540 (Edin.: Pentland), or Hermann. *H'dbch d. Physiol.* (Leipzig: Vogel), 1879, II, **1**, 171.

‡ Rosenberg. *J. Physiol.*, 1935, **84**, 50.

§ Hill. *J. Physiol.*, 1935, **83**, 30P.

The accommodation constant may be used to explain many facts in connection with nerve physiology. Thus, with a weak constant or slowly rising current, excitation will not occur until V reaches and passes U_X. Again, the anode cannot excite at make because the anodic V cannot catch up with U, but at break the V returns faster to V_0 than U does to U_0, so that excitation takes place.

Electrotonic changes also occur in other excitable tissues, such as muscle,† but these cannot be considered here.

For Reference

Ebbecke (1933). Zur Lehre vom Elektrotonus. *Ergebn. d. Physiol.*, **35**, 757.

Kugelberg (1944). Accommodation in Human Nerves. *Acta Physiol. Scand.*, **8.** Suppl. xxiv., pp. 1–104.

Lorente De Nó (1946). Correlation of Nerve Activity with Polarisation Phenomena. *Harvey Lects.*, **42,** 1946–47, 43.

Schmitt and **Bear** (1939). The Ultrastructure of the Nerve. Axon Sheath. *Biol. Rev.*, **14,** 27.

Skoglund (1945). The Electrotonic Changes of Excitability in Nerve and their Relation to Accommodation. *Kungl. Svensk. Vetenskapsakad. Handl.*, **21,** 5.

ELECTRICAL STIMULI AS APPLIED TO HUMAN NERVES

When we attempt to apply the results gained on frog's nerves to man, we are met at once by the difficulty that we cannot apply stimuli to human nerves directly. So usually *unipolar* excitation is used, a small or stigmatic electrode, either anode or cathode, being applied to the skin over the nerve to be stimulated, and the other, a large one, to some indifferent area, such as the back. Under these circumstances the current is concentrated as it leaves the stigmatic electrode, and diffuses widely in the body, seeking the lines of least resistance. Thus it is impossible to get pure anodic or cathodic effects. If the anode be applied over the nerve, the current enters by a series of points (the polar zone), and leave by a second series (the peripolar zone). The polar zone will thus be in the condition of anelectrotonus, and the peripolar zone in that of catelectrotonus (Fig. 113). The currents however, will be more concentrated at the polar than at the peripolar zone, and so the former effect will predominate. These restrictions in the application of the current cause slight apparent irregularities in the "law of contraction" as tested on man.

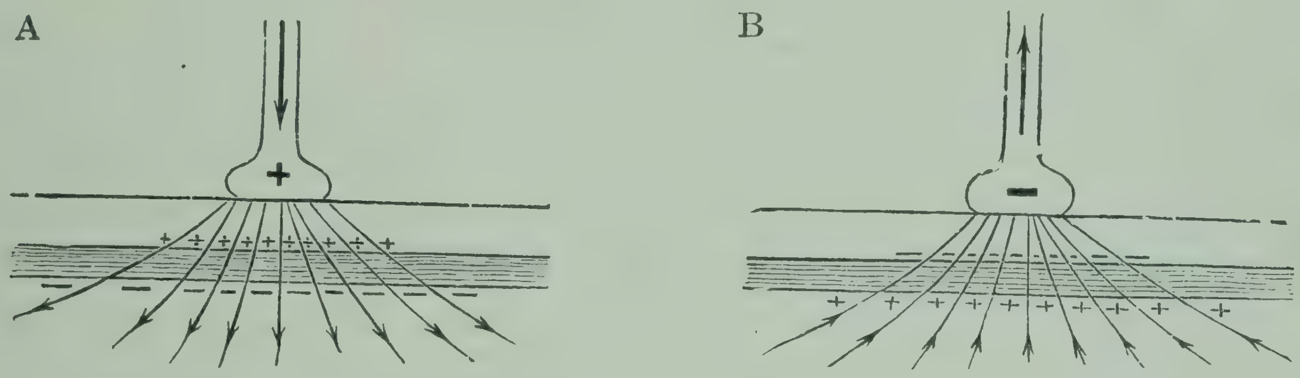


Fig. 113. Electrodes applied to the Skin over a Nerve Trunk.

In A the polar area is anelectrotonic and the peripolar catelectrotonic. The former condition preponderates, since the current here is more concentrated. In B the conditions are reversed, the polar zone corresponding in this case to the cathode. (Waller.)

In stimulating the nerves of man for the purpose of determining the conditions of the different muscles, we may use either faradic or galvanic stimulation. The electrodes are covered with chamois leather moistened with salt solution in order to diminish the resistance of the skin. When it is desired to stimulate any given muscle, the stigmatic electrode is brought as nearly as possible over the spot where the muscle receives its motor nerve. These *motor points* have been mapped out, and reference is generally made to a diagram in determining the point for any given muscle. By reversing the current the stimulating lectrode may be made either anode or cathode. It is found that stimulation occurs most asily on closure of the current with the stimulating electrode the cathode; with the greatest difficulty when the current is broken and the stimulating electrode is the cathode.* These different contractions are generally represented by capital letters, and in order of increasing difficulty of elicitation we have :—

---

* Hill and Solandt. *J. Physiol.*, 1935, **84,** 1P.

† Schaefer and Schölmerich. *Pflüger's Arch.*, 1938, **241,** 310.

CCC = cathodal closing contraction.
ACC = anodal closing contraction.
AOC = anodal opening contraction.
COC = cathodal opening contraction.

The problem of the determination of the chronaxies of the human nerves or muscles has been investigated by Bourguignon,† who made the important observation that different muscles (with their nerves) have different chronaxies, and further, that in the same segment of a limb, muscles of similar function have the same chronaxie. Thus when the chronaxie of flexor and extensor muscles is compared it is found that the flexors have a chronaxie only half as great as the extensors. This agrees with the relative duration of the contractions, which Baird and Fulton ‡ find to be about 30 msec. and 40 msec. respectively.

When the motor nerve to a muscle has undergone degeneration, the muscle also begins to degenerate, and we find certain alterations in its chronaxie. For some time after section of the motor nerve the chronaxie of nerve and muscle remains unaltered ; this period may be up to a week or ten days in the sciatic-gastrocnemius of the frog, and at the end of this time the nerve has ceased to function. The chronaxie of the muscle then steadily increases, and at the end of three months is up to five times its normal value. In warm-blooded animals and in man the change takes place much more rapidly.

The muscle after degeneration of its nerve gives a slow contraction and is frequently of such high chronaxie as to be inexcitable to the brief faradic shock (reaction of degeneration). In the second place, qualitative alterations in excitability may be present, so that ACC may be obtained with a smaller current than CCC. Changes in the chronaxie are, however, detectable long before this stage is reached.

THE NATURE OF THE NERVE IMPULSE

Under this heading we have two questions to discuss : (*a*) the nature of the change aroused at the stimulated spot in the fibre, and (*b*) the mode of propagation of the excitatory change away from the stimulated spot and along the fibre.

The nature of the electrical potential changes, the all-or-none character of the impulse, together with the normal absence of decrement in the excitatory process, have been taken as an indication that the plasma membrane of the axon is the seat of energy changes which may be let loose under the influence of a stimulus.

It is not known how some types of stimuli, *e.g.* mechanical, act, but in some way they alter the local state of polarisation of the plasma membrane, causing either depolarisation, or a reversal of polarisation. For this reason electrical stimulation by a cathode is far the most efficient, since in this way an immediate negative charge is imparted to the outer membrane surface. As a result of a structural change which this induces in the membrane, this now becomes freely permeable to ions to which it was previously impermeable. If the applied cathode potential is a feeble one, and small in extent, the result is merely a drop in the resistance of the membrane and a local non-propagated response : but if it is powerful and extensive enough to reverse the polarity of the membrane over a sufficient area, then a

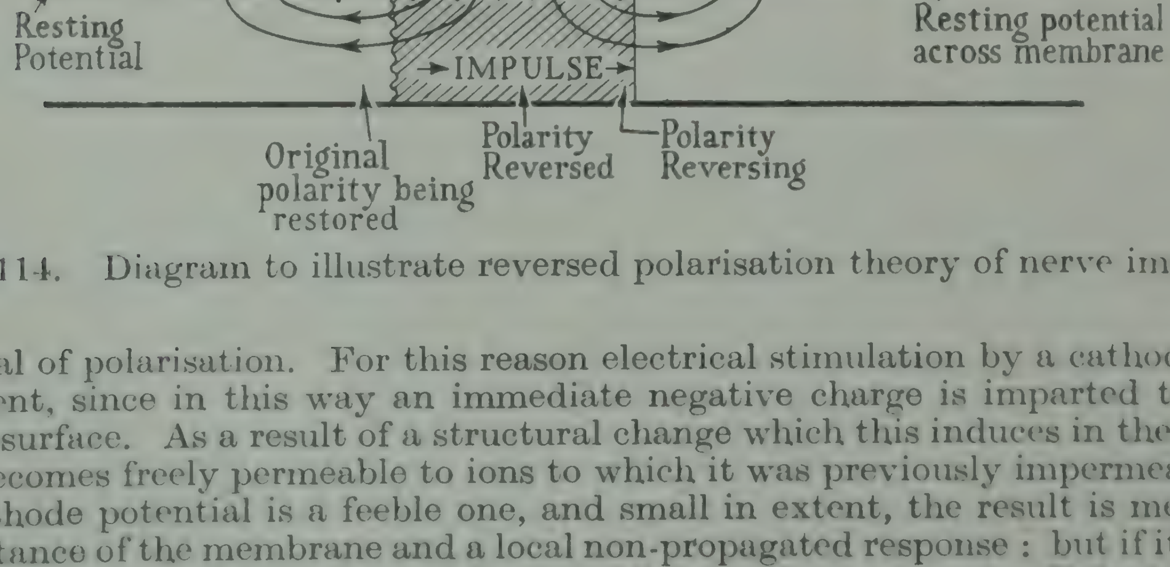


FIG. 114. Diagram to illustrate reversed polarisation theory of nerve impulse.

---

* DE WATTEVILLE. L'electrotonus des nerfs chez l'homme. 1883. (London : De Ranken.)

† BOURGUIGNON. La chronaxie chez l'homme. 1923. (Paris : Masson.)

‡ BAIRD and FULTON. *Amer. J. Physiol.*, 1927, **81,** 462.

propagated impulse is started off. Study of the excitability changes beyond a block prove that the propagation is electrical in character, current flow from resting to active membrane along the outside, and in the return direction on the inside of the membrane, as shown in Fig. 114, acting in the same sense, and producing on nearby resting membrane the same result, as an applied cathode. The process thus spreads from section to section along the nerve fibre, as originally postulated by Herman in 1879.* At the portions occupied by the impulse, the polarity is reversed and the properties of the membrane temporarily altered, and no stimulus, however strong, can produce any further effect. Hence the refractory period.

These changes do not involve any considerable chemical change or heat production, the energy for the impulse being stored up in the membrane as polarisation charges. It is probable that the surface membrane, largely composed of lipoid, separating the axis cylinder from its surroundings, is the site of the propagated disturbance. If we suppose that this is only two molecules thick,† it is evident that the amount of energy involved would be very small, as the volume concerned would be only a very small portion of that of the whole fibre. Thus the active element in a nerve fibre would be a bimolecular film maintained in a polarised state at the expense of the oxidative reactions which normally go on slowly during life. These reactions, and hence the chief energy expenditure, are concerned with restoring the state of polarisation after the impulse has passed. Hinder the oxidation and the film remains depolarised, so that the impulse can no longer be transmitted.

After the impulse has passed along the fibre, this is repolarised in the resting direction : the process takes time, during which the fibre is in a relatively refractory state.

The depolarisation during activity is accompanied, or perhaps caused, by an increase in the permeability of the nerve membranes. In consequence of this, certain substances, notably potassium, pass out of the fibre into the surroundings and Na enters it. We may suppose that it is this passage of potassium ions through membranes previously able to restrain them from wandering, which constitutes the essential cause of the depolarisation. Similar phenomena occur in muscle, and it is perhaps worthy of note that both tissues are readily rendered inexcitable when dilute solutions of potassium salts are applied to their surfaces.

Acetyl choline is released from cholinergic fibres and sympathin from adrenergic ones during the passage of an impulse ; thiamine is also said to be set free. Whether, as some suppose, these are the cause, or as others guess, the effects of the electrical depolarisation has yet to be decided.

Lillie's Model. R. Lillie has drawn attention to the striking analogy between the nerve impulse and the change which can be propagated along a "passive" iron wire. If iron wire be immersed in nitric acid of a specific gravity exceeding 1·2, it reacts only momentarily and then enters into a passive state in which no further solution occurs, a thin film of oxide being developed and maintained on its surface. Yet this seemingly stable system is highly reactive. Activation may be brought about by electrical, chemical or mechanical means. For example, if the wire is touched by a piece of zinc, a local electrical circuit is produced and it becomes activated at that spot, and turns black and effervesces ; moreover, the activation thus initiated spreads as a wave of definite velocity over the whole surface of the wire, lasts for a few seconds, and disappears, the iron becoming passive again. But now the freshly passivated iron is resistant to re-activation for a minute or so (refractory period), after which the sensitivity to activation returns again. Mechanical stimulation, summation, electrotonus, polar stimulation on electrical activation, etc., can all be imitated on this fascinating model, as can also some of the facts of the time-relations of electric currents, such as the relative inefficacy of slowly rising currents to activate as compared with currents of abrupt onset. Thus the iron wire in these particular conditions transmits a self-propagating electrochemical change, limited to its surface film and in many respects analogous to the propagated impulse in a nerve.

TRANSMISSION OF EXCITATORY STATE FROM NERVE TO EFFECTOR TISSUE

The excitatory process travelling down a motor nerve is transmitted to the muscle by the intermediation of the motor end-plate. A similar break seems to occur in the central nervous system wherever an impulse is propagated from the axon process of one nerve cell to the body or dendrites of another nerve cell, and the area of contiguity, where an impulse passes from one neuron to another, is spoken of as a *synapse*. The presence of the end-plate between muscle and nerve imposes certain new conditions on the conduction of the excitatory state. One of the most important of these lies in the fact that the conduction across the end-plate, like that across the synapse of the central nervous system, is *irreciprocal*. An excitatory process started in the nerve is transferred easily to the muscle. But an excitatory process started in the muscle does not extend back to the nerve. This fact may be shown on the frog's sartorius. If the upper (pelvic) end of the muscle be split, as in Fig. 115, a stimulus applied to the nerve-free end of one of the limbs, *e.g.* at A, causes a contraction, with accompanying electrical changes, in the corresponding half of the muscle, but no impulses, *i.e.* no electric changes, in the motor nerve.

---

* Hermann. *H'dbch. d. Physiol.*, 1879, II, **1**, 193.

† Danielli. *J. Physiol.*, 1941, **100**, 117.

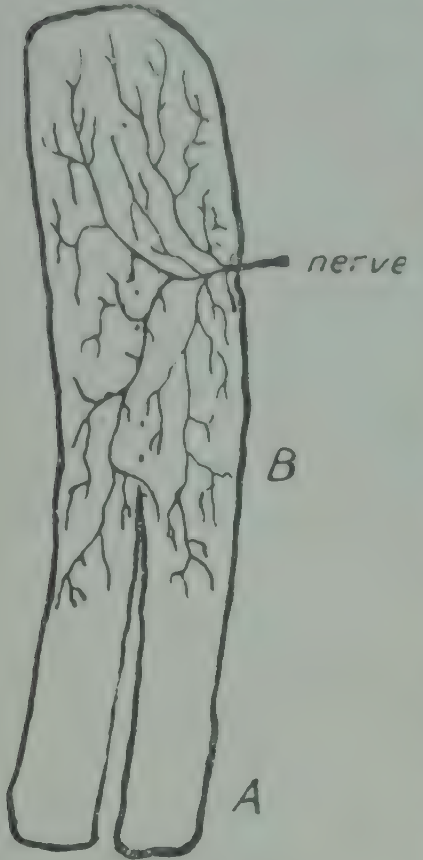


FIG. 115.

The transmission of the excitatory condition across the end-plate from nerve to muscle occupies time. If we compare the latent period of a muscle stimulated directly with its latent period when excited through the nerve, we find that there is an increased period of delay in the latter which is not wholly accounted for by the time taken for the impulse to travel from the stimulated nerve point to the muscle. The extra delay is due partly to the fact that the terminal branches of the nerve fibres supplying any motor unit are much finer than the axons from which they spring; they therefore probably conduct more slowly. But there is further delay due to processes occurring in the end-plate. The total delay in passing from the coarser fibres to the muscle is probably about 3 to 4 msec. for the frog's gastrocnemius, but for rapid mammalian (*e.g.* eye) muscles, it is only 0·7 msec.*

The spread of excitation from the end-plate seems to be the weakest point in the neuro-muscular chain. We have already seen that, when the nerve of a nerve-muscle preparation is stimulated repeatedly, the muscle very soon shows signs of fatigue, and that the seat of this fatigue is not in the nerve nor in the muscle, but at the junction between the two. Similarly, curare quenches the spread from end-plate to muscle.

The classical electrophysiological view is that the excitation of muscle through nerve depends on the electrotonic potential changes occurring at the nerve ending. The occurrence of such *end-plate potentials* (e.p.p.) can best be demonstrated in well-curarised muscle; when the motor nerve, *e.g.* to a strip of curarised cat soleus, is stimulated, a potential is set up in the part rich in end-plates.† It is due to a local depolarisation, reaches a peak in 2·5 msec., decays exponentially, and in curarised muscle (or during the refractory state) suffers a decrement of at least 50 p.c./mm., so that it never travels far. In partial curarisation a second volley falling a few msec. after a first produces a greatly increased e.p.p., followed by depression of several seconds.‡

Comparison with a Partial Block. In many respects the conduction of impulses along a partially blocked nerve resembles that across the neuromuscular junction or across a synapse. The properties of a nerve block were first studied from this standpoint by Wedensky.§ Heat, cold, galvanic polarising currents or various chemical agents, *e.g.* cocaine or alcohol have been used to cause the partial block. A partly blocked nerve will conduct impulses at low frequency, but not at high frequency—a phenomenon called Wedensky inhibition. In the blocked region there is a raised threshold of excitation, but beyond the block the threshold is lowered. This difference is accentuated if impulses at a rapid rate are sent into the region and fail to pass through it. When this happens, there is for every impulse arrested at the block a negative electrotonic potential below the block.* When the summated electrotonic potential reaches a critical value, slightly above that of a spike potential, an impulse is generated, thus " leaping " the block. It is believed that a similar state of affairs obtains at neuromuscular and synaptic junctions.†

---

* LORENTE DE NÓ. *Amer. J. Physiol.*, 1935, **111**, 272.

† ECCLES and O'CONNOR. *J. Physiol.*, 1941, **100**, 317.

‡ BROWN and HARVEY. *J. Physiol.*, 1941, **99**, 379.

§ WEDENSKY. *Pflüger's Arch.*, 1903, **100**, 1.

Humoral Transmission. There is now reason to believe, however, that the excitatory state set up in the muscle is associated with, and perhaps due to, the liberation, at or near the nerve ending, of the potent chemical substance acetyl choline, which is liberated at the end-plates by the arrival of the nervous impulse, rapidly exerts its action in producing contraction, and then is rapidly destroyed. The denervated muscle, as mentioned on p. 136, responds to a dose of acetyl choline by passing into contracture. If small doses (2–20μg.) of acetyl choline are injected direct into the arteries to a normal muscle, quick contraction of the muscle results, and if the muscle is previously eserinised, the response to acetyl choline or to nerve stimulation is increased : curarine abolishes the response to acetyl choline, but atropine does not. Curarine probably acts by combining with some receptor group in the end-plate material and shielding it from combination with acetyl choline.‡

The end-plates contain the enzyme choline esterase in a concentration many thousand times as high as in muscle fibres alone. Hence the acetyl choline could be destroyed with the necessary speed during the refractory period.§ The release, and subsequent destruction, of acetyl choline at its sites of operation are estimated to take place in time intervals of the same order of magnitude as those of the electrical changes of the spike potential.

In the disease myasthenia gravis, the subject shows extreme muscular weakness and fatigability. This is believed to be due to either excessive activity of cholinesterase or failure of formation of acetyl choline at the motor end-plates. The patients are restored to muscular strength in spectacular but transitory manner by injection of prostigmine, a synthetic analogue of eserine with similar anticholinesterase activity. Other anticholinesterase agents have also been tried, but these often have general toxic effects.

It would seem, therefore, that liberation of acetyl choline at the nerve endings may be a factor in the passing on of excitation from nerve to muscle, and it will be shown later (Chapter XXIII) that the transmission of excitation at other points of junction in other tissues is probably also effected by similar chemical transmitters, and it has even been suggested that the substance is responsible for the brief and reversible increases in ionic permeability which lie at the root of the depolarisation manifesting itself as the electric potential change in nerve fibres and other tissues. Very possibly connected with this, movements of potassium and calcium ions are also concerned, either primarily or secondarily, in the process of transmission.

For Reference

ADRIAN (1930). The Mechanism of the Sense Organs. *Physiol. Rev.*, **10**, 336.

BROWN (1937). Transmission at Nerve Endings by Acetyl Choline. *Physiol. Rev.*, **17**, 485.

DAVIS (1926). The Conduction of the Nerve Impulse. *Physiol. Rev.*, **6**, 547.

ECCLES (1937). Synaptic and Neuro-muscular Transmissions. *Physiol. Rev.*, **17**, 538.

GERARD (1931). Nerve Conduction in Relation to Nerve Structure. *Quart. Rev. Biol.*, **6**, 59.

LILLIE (1922). Transmission of Physiological Influence in Protoplasmic Systems, especially Nerve. *Physiol. Rev.*, **2**, 1.

LILLIE (1936). The Passive Iron Wire Model of Protoplasmic and Nervous Transmission and its Physiological Analogies. *Biol. Rev.*, **11**, 181.

LORENTE DE NÓ (1947). A Study of Nerve Physiology. *Stud. Rockefeller Inst. for Med. Res.*, Vols. 131, 132. (New York.)

NACHMANSOHN *et al.* (1946). The Physico-chemical Mechanism of Nerve Activity. *Ann. N.Y. Acad. Sci.*, **47**, 375–602.

SYMPOSIUM (1948). The Physiology of Acetyl Choline. *Bull. Johns Hopkins Hosp.*, **83**, 463.

---

* HODGKIN. *J. Physiol.*, 1937, **90**, 183, 211.

† LORENTE DE NÓ. *J. Neurophysiol.*, 1939, **2**, 402.

‡ BROWN, DALE and FELDBERG. *J. Physiol.*, 1936, **87**, 394.

§ MARNAY and NACHMANSOHN. *J. Physiol.*, 1938, **92**, 37.

BOOK III

CENTRALIZED CONTROL AND CO-ORDINATION—THE CENTRAL NERVOUS SYSTEM

Books for General Reference

CREED, DENNY-BROWN, ECCLES, LIDDELL and SHERRINGTON (1932). Reflex Activity of the Spinal Cord. (Oxford : Clarendon Press.)

FULTON (1943). Physiology of the Central Nervous System. (Oxford University Press.)

HERRICK (1934). An Introduction to Neurology. (London : W. B. Saunders & Co.)

HOLMES (1946). Introduction to Clinical Neurology. (Edinb.-Livingstone).

MAGNUS (1924). Körperstellung. (Berlin : Springer.)

RANSON (1947). The Anatomy of the Nervous System from the Standpoint of Development and Functions. (Philadelphia : Saunders.)

SHERRINGTON (1906). The Integrative Action of the Nervous System. (London : Constable.)

SHERRINGTON (1939). Selected Writings. Ed. by Denny-Brown. (London : Hamish Hamilton.)

WALSHE (1948). Critical Studies in Neurology. (Edinburgh : Livingstone Ltd.)

HISTORICAL NOTE

Scholarly curiosity has unearthed many early philosophic speculations regarding the functions of the nervous system, but space will not allow of their being mentioned here. It is true that DESCARTES (1596–1650), one of the later of these philosophers, evolved speculations (publ. 1662) in which considerable foresight was subsequently discerned, or supposed ; but the lack of experimental, or even of anatomical, basis for these theories, robbed them of much of their scientific value, as STENSEN, indeed, pointed out.

The essential anatomical foundations began to be laid by THOMAS WILLIS (1659) and by STENSEN (1662), and have been continued by many since ; the preliminary hardening of tissue (REIL, c. 1800) ; the method of serial sections (STILLING), and of staining (GERLACH, 1858), of dissociation by forceps (MEYNERT), and the introduction of the WEIGERT method (1884) all played an important part in these increasingly detailed anatomical studies. It is more difficult to say at what date instructive and well-directed experiments were initiated, but, apart from those of GALEN, which for centuries were unique and disregarded pointers, perhaps some observations on decapitate frogs by Johann BOHN in 1686, and the ablation of the cerebrum and cerebellum by du VERNEY in 1697 may be accepted as early beginnings.

HALLER (1757) recognised that nerves were essential to both feeling and movement ; and conveyed messages to and from the brain, which he thought could also feel.

The first outline of the principle of reflex action, and the demonstration of the dependence of each reflex on the integrity of a particular part of the central nervous system, was given by ROBERT WHYTT (1768) : the conception of a reflex *arc*, however, had to wait until a clear distinction had been made between afferent and efferent nerves. This emerged with the demonstration of the functions of the ventral and dorsal spinal nerves by BELL (1811) and MAGENDIE (1822). The logical sequel, the conception of the reflex arc, was introduced by MARSHALL HALL (1833 ; 1850), who also showed the reflex nature of sphincter action, and the actions of chloroform and strychnine on spinal reflexes.

Although GALEN, in the 2nd Century had demonstrated that the paths for both movement and sensation passed along the spinal cord, it was not until the fifth to the seventh decades of the 19th Century that the approximate functions of the dorsal, ventral and lateral columns of the cord were discovered by the methods of section and stimulation (LONGET, TÜRCK, VULPIAN, SCHIFF, BROWN-SEQUARD, WOROSCHILOFF and others). The important fact that the fibres of the dorsal roots do not arise from cells of the grey matter was discovered by GERLACH (1871). The method of nerve degeneration discovered by WALLER (1850) not only allowed of the following out of nerve tracts, but supplied the important theoretical conception of the dependence of the axon, discovered by WAGNER (1851), upon the cell. Out of the realisation of this, and as a result of histological work by MARCHI, HIS and others, there grew the neurone theory (WALDEYER, 1891), which was supported and confirmed by the work of Golgi, Cajal and others, and which, together with the conception of the synapse by Foster (1897), has contributed so largely to the establishment of present-day views of the nervous system.

Not long after the promulgation of the idea of reflex action there appeared indications that conduction in the C.N.S. differed in many respects from that along nerves. Helmholtz (1854) showed that reflex time might be 12 times as slow as nerve conduction, and Exner (1873) that the transmission through the centres was generally slow: also Du Bois Reymond (1859) gave reasons for supposing that the rate of emission of the messages along motor nerves from the centres differed from their frequency when the motor nerves were stimulated directly.

Early speculators had often suggested that the different parts of the C.N.S. had different functions, but their enquiries were largely concerned with the search for the seat of the soul, which was variously located, *e.g.*, in the pineal body, the medulla, or even diffused over the whole body: Willis supposed that the corpora striata were the seat of sensation and movement. It was not until the early decades of the 19th Century that experimental ablations were practised with noteworthy results. Rolando (1809) showed that voluntary movement was lost after removal of the cerebral hemispheres, and Desmoulins (1825), Flourens (1842) and Goltz (1868) studied the various effects of removal of parts of the cerebral cortex, but without any definite ideas of localisation of function emerging. In the meantime there had been introduced, by Gall and Spurzheim (1812), the views on localisation of cerebral function which subsequently became known as Phrenology, and were soon quite discredited. The discovery, however, by Broca (1861) of an area in the brain the integrity of which was definitely essential for speech, was the first clear indication that there might be some localisation of function in the cerebrum, and this was further supported by the anatomical investigations of Meynert, the clinical observations on epileptiform convulsions by Hughlings Jackson (1868), and especially by the experiments of Fritsch and Hitzig in 1871, who showed that galvanic excitation of certain areas of the cortex resulted in precise movements of the limbs. The substitution of faradic for galvanic excitation by Ferrier, and his extensive exploration of the cortex of the ape (1874) firmly established the notion. Of sensory areas, the visual was studied by Ferrier, and its approximate location was fixed by Munk (1878). The motor areas in man were forecast in 1873 by Hitzig, the characteristic large cells of the area being discovered by Betz in 1874, and further confirmation of the functions of this area provided by Ferrier in 1879, and by Sherrington (1895–1917). Following the observations of Baillarger (1840), v. Kölliker (1885) and Meynert (1868) that lamination varied from place to place in the cortex, the mapping of the cortex into areas of characteristic structure and lamination was actively pursued (Bolton, Brodman, Campbell and others) until early in the present century. The fact remained, however, and still remains, that, apart from certain special areas to which specific functions can be attributed, large areas of cortex can be removed without any very serious effects, as was evidenced by a series of clinical cases ranging from that of Berenger de Carpi (1715) to the famous "American crowbar accident" of Bigelow (1850). The idea of a functional integration of the cerebral cortex, adumbrated by Foster (1897) is not necessarily in opposition to the view that there is localisation of certain executive zones of the cortex. The experiments of Pavlov (1898–1935) on the conditioned reflexes have perhaps provided reasons for such a compromise in outlook.

The analysis of the phenomena of reflex action, as exemplified by the spinal reflexes, forms the hard core of our knowledge of the physiology of the central nervous system. Apart from observations on irradiation of reflexes (Pflüger, 1853), on the fact of the possibility of central inhibition (Setschenov, 1863) and of facilitation (Exner, 1882), little real progress was made in the study of the spinal reflexes until the closing decades of the 19th Century, when the work of Sherrington began: he introduced the method of myographic analysis for the quantitative study of motor reflexes, and his work, outlined in "The Integrative Action of the Nervous System" (1906), forms the basis of our present-day knowledge of the reflex as an integrating mechanism of high precision, and threw a flood of light on the working of the entire central nervous system. The elegant studies of Magnus on postural reflexes provided a valuable complement to that work.

The location of various "centres" in the brain stem was sporadic throughout the 19th Century. The medullary respiratory centre was identified by Legallois in 1812; the corpora quadrigemina as a centre for light reflexes by Flourens in 1847; the medullary vasomotor centre by Ousjannikov in 1871, and so on. Views on the levels of integration of nervous functions took shape about 1884 largely under the influence of Hughlings Jackson.

Much labour has been expended in investigations on the cerebellum; the effects of its ablation were studied by Flourens (1822), Magendie (1824), and many others since, some of the most profitable studies being those of Luciani about 1890. It was first regarded as the seat of muscle sense by Lussana and Morganti in 1853.

The discovery of the autonomic system has been of slow growth. Galen described the vago-sympathetic trunk, the sympathetic chain, ganglia, rami, and splanchnics as branches of his sixth pair of nerves, and it was not until 1552 that Eustachio recognised the separate nature of the vagus. He described the sympathetic, however, as arising from the cranium. Willis, in 1644, gave the name of "intercostal nerve" to the sympathetic chain, and regarded it and the vagus as responsible for involuntary actions, and for "sympathetic" relations between various parts of the body, a conception which was widened considerably by Whytt (1751–1756).

It was Pourfour de Petit (1727) who first showed that the sympathetic did not arise from the cerebrum. The ganglia of the sympathetic were considered by Winslow (1732) and by Johnston (1764) to be independent centres, and it was not until the 19th Century that their real nature was understood. The independence of the ganglia was further stressed by Bichat (1801); he saw that the sympathetic was not as other nerves, but was a chain of ganglia, and he separated the nervous control of somatic from that of visceral functions, which were involuntary, and controlled such things as digestion, circulation, etc.

The cell bodies of the ganglia, and the white and grey fibres were distinguished by Ehrenberg (1833), and Valentin (1836) first recognised the white and grey rami, and their relation to the cord, while the first accurate description of the characteristic non-medullated fibres was given by Remak in 1838.

The vaso-motor functions of the sympathetic were simultaneously recognised by Claude Bernard and by Brown Sequard in 1852, and the ventral roots recognised as the origin of the sympathetic.

Embryological proof that the sympathetic arises from the spinal cord was given by Schenk and Birdsall in 1878, and in 1885 the important work of Gaskell began. He showed the relation between the white and grey rami, mapped out their anatomical distribution, and gave to the system the name of the involuntary nervous system: he divided this into bulbar, thoracicolumbar and sacral outflows, and noted an underlying antagonism between those nerves of the anabolic type, *e.g.* the vagus, and the katabolic sympathetic. The further understanding of the anatomical arrangement, and the distinction between pre- and post-ganglionic fibres (1893) was based on the nicotine method introduced by Langley and Dickinson in 1889. The name "autonomic" was introduced by Langley in 1898, and the "parasympathetic" and sympathetic were separately designated by him in 1905. This distinction was largely based on the action of drugs, especially of that of adrenaline, which Langley showed in 1901 closely resembled sympathetic stimulation in its effects. Later, the humoral theory of nervous action was introduced, chiefly as the result of the work of Loewi, Dale, Cannon, Magnus and others, and the distinction of cholinergic from adrenergic nerves introduced by Dale in 1934.

Modern work, employing as powerful tools the micro-electrode, valve amplification and the cathode ray oscillograph, is rapidly developing research in the C.N.S. in a variety of directions, and many important discoveries have already been achieved by these means.

CHAPTER XIII

GENERAL FEATURES OF A CENTRAL NERVOUS SYSTEM

THE central nervous system enables the individual to respond in a co-ordinated manner to changes in its environment. In the course of its evolutionary development the cells and their processes, the nerve fibres, of which it is essentially composed, have been endowed in an exaggerated form with the properties of excitability and conductivity inherent to lesser degrees in all living protoplasm. Being thus highly excitable and able to conduct the excitatory state with speed and precision for considerable distances, and often to distribute it over wide areas, it is able under appropriate conditions to put remote parts of the body in temporary functional connection with one another and by means of these connections to give close correlation between the actions of such parts.

A nervous system is functionally divisible into three series of parts, the *afferent*, the *central* and the *efferent*. The afferent portions consist of those nervous structures which normally originate and conduct nervous impulses towards the centres, and the efferent portions of those conducting outwards from the centres.

The afferent portions are connected with appropriate sensory receptors and serve to convey nerve impulses aroused by stimuli, *e.g.* visual, touch, etc., from the outer world (exteroceptive stimuli), or stimuli arising in the body itself, *e.g.* in viscera or muscles (interoceptive stimuli), to the nerve centres.

The most important stimuli which reach us from the outer world are those of light and sound and to a lesser degree (in man) of smell, since by these we obtain knowledge and warning of things at a distance from us. Taste, by enabling us to make judgments as to edibility is next in importance. These stimuli are all received by organs placed in the head, which therefore has the prior position in the body. Touch is more generalised and requires actual contact with the sensory surface. It is confined to the surface of the body, the deeper structures of which are totally insensitive to touch.

Of the interoceptive stimuli, important are those (proprioceptors) that arise in the muscles, joints, tendons, etc., which give us information as to the position of our limbs, etc., those from the labyrinth which tell us of the position of the head in space, or of its movement, and those from the various internal organs. Many of the interoceptive stimuli arouse no conscious sensations.

The nerve centres receive these impulses, some of which enter consciousness and produce an awareness of the world around us, and of the position and state of our own bodies. It is obvious that our knowledge of the universe, including that of our own bodies, is solely dependent upon, and limited by, these impressions. Many nervous impulses arriving at the centres do not enter consciousness at all, but nevertheless they play an important part in maintaining or regulating the many processes which are constantly proceeding, unknown to us, in the body.

The efferent parts of the system conduct impulses outwards from the centres, and these impulses represent the responses of the centres to the afferent impulses which they have received. When the efferent impulses so generated in, or reflected outwards by, the central portions arrive at their destinations in the so-called *effector organs*, *e.g.* muscle, gland, etc., these are thrown into action and this action represents the response mediated by the central portions as appropriate to meet the particular situation indicated by the afferent stimulus.

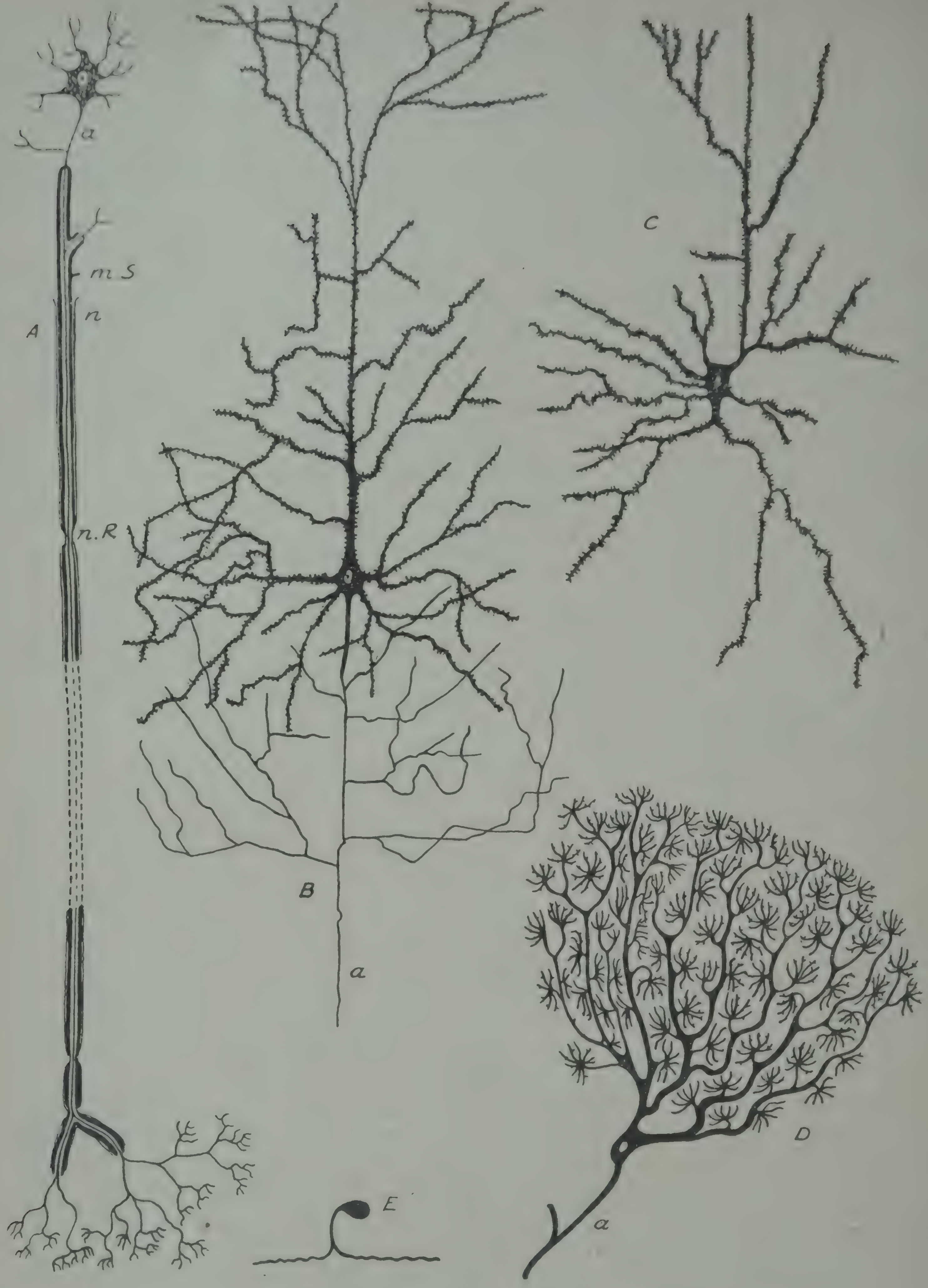


FIG. 116. Characteristic Types of Neurones (diagrammatic).

A Multipolar motor neurone of spinal cord.
B Pyramidal cell of cerebral cortex.
C Golgi Type II cell of cerebral cortex.
D Purkinje cell of cerebellum.
E Bipolar cell of spinal ganglion.
a The axon, is indicated in *A*, *B* and *D*.
n Neurolemma ; *m.S.* medullary sheath ; *n.R.* node of Ranvier.

Two main types of cells go to build up the nervous system, viz. *neurones* and *neuroglia* cells. The latter are supporting cells and are of no immediate physiological interest.

THE NEURONE is the name given to a nerve cell with all its branches. According to the neurone theory, now generally accepted, the neurone is the functional unit of the central nervous system. There are many kinds of neurones, classified variously according to situation, shape, size, number and length of branches, etc., but every neurone consists of two essential parts, viz. the *soma*, neurocyton or nerve-cell body, and second, its branches, which are nerve fibres (Fig. 116).

The branches of each neurone may vary from one to several in number, but one, and only one, of them is distinguished as the axis cylinder process or *axon*, the others being called dendrites or *dendrons*. The axon carries nerve impulses away from the soma, and the dendrons convey them towards the soma.

The neurocytons, or nerve cell-bodies, are found in the grey matter of the central nervous axis and in the various ganglia which lie outside it, but not in the white matter, which consists of nerve fibres only. Every nerve fibre, however, is a branch, either an axon or a dendron, of a nerve cell.

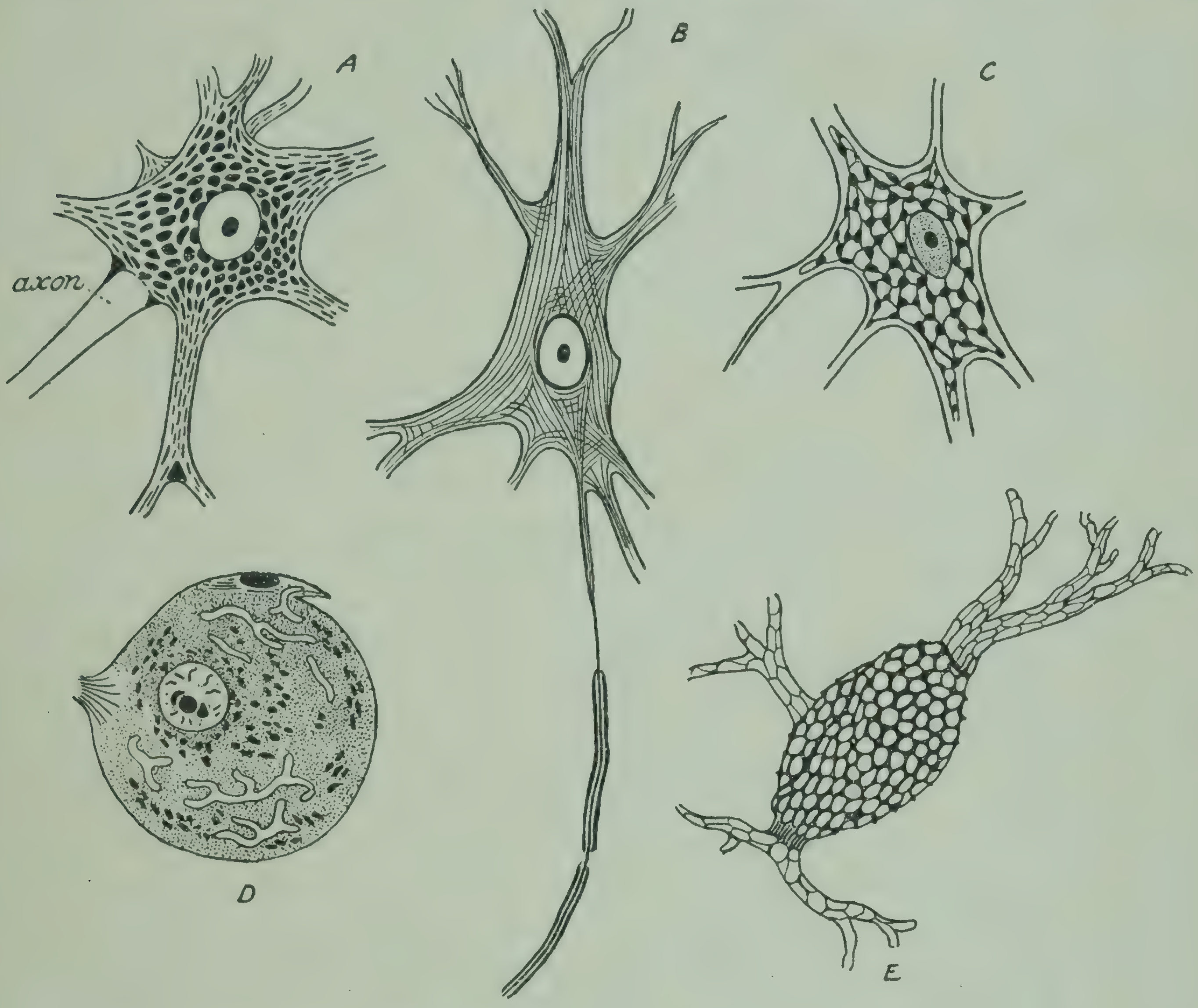


FIG. 117. Illustrating Structure of the Neurocyton.

A, Nissl granules in multipolar cell; *B*, neurofibrils in multipolar cell, and addition of medullary sheath to axon; *C*, Golgi endo-reticulum in multipolar cell; *D*, Holmgren's canaliculi, in ganglion cell; *E*, Golgi superficial network in multipolar cell.

The neurocyton, whatever its type, always exhibits certain characteristics of structure on examination by special histological fixing and staining methods.* Although these characters have often been claimed to be produced by the methods employed for fixation and staining, their invariable presence in cells under certain physiological conditions, and their absence or modification in other circumstances, give them an interest and a claim to further study (Fig. 117). There are, for instance, the Nissl's bodies,† and the Golgi networks, neither of which are visible on ordinary inspection in the living cell-bodies. The Nissl granules are, however, easily seen by photography with ultraviolet light.‡

The Nissl bodies are spindle-shaped masses of basophile material, apparently consisting of a nucleoprotein containing iron. They are of a size, and are arranged in a manner, characteristic for each type of neurone, and are disintegrated in fatigue, by certain poisons, and after section of the axon of the neurone, this process being called *chromatolysis*. Nissl bodies are present in all the hillocks from which dendrites arise, but are absent from the axon-hillock for some unknown reason.

---

* GREENFIELD, *J. Neurol. & Psych.*, 1938, **1**, 306.

† MOTT. *B.M.J.*, 1912, *ii*, 780.

‡ WEIMANN. *Z. ges. Neurol. Psych.*, 1925, **98**, 347.

The Golgi networks are of two kinds, an endocellular and a superficial. Their functions are unknown, but following section of the axon the endocellular reticulum is displaced to the periphery of the cell and often fragmented. The fine channels called Holmgren canaliculi are of doubtful significance, and are probably artefacts.

The neurofibrils seen in the cell-body on special fixation and staining form intricate networks continuous with the neurofibrils demonstrable in the axon and dendrons; their properties vary with the physiological state of the cells, and although some consider them to be artefacts, others claim to have demonstrated them in the living cells and fibres.

Each neurocyton has its nucleus, but no centrosome, and it is perhaps for this reason that adult nerve cells are incapable of division. Mitochondria are present in the cytoplasm, where they are visible in the living state.

Many neurones are entirely confined to the grey matter of the central nervous system, and these as a rule have only short axons. They are called Golgi type II. cells (Fig. 116, C). Those neurones that send their axons into the white matter, or into peripheral nerves are called Golgi type I. cells.

The dendrons of a nerve cell, and those parts of its axon which lie within the grey matter of the central nervous system, are composed of neuroplasm and neurofibrils, with no demonstrable cell wall or sheath, but on emerging into the white matter a covering is usually added to the axon, viz. the medullary sheath. To those axons that lie outside the central nervous system a second covering, the neurolemma is added.

Terminations of Nerve Fibres. The medullary sheath and the neurolemma are present throughout the peripheral nerves until near their terminations (in efferent fibres) or their origins (in afferent fibres). As these terminations (or origins) are approached the axons branch freely, at nodes of Ranvier, so that several, or many, terminal structures are connected with one axon (Fig. 116, A). Motor fibres for instance usually form about 100 branches, each one supplying one muscle fibre. The name *motor unit* is given to the complete structure consisting of motor neurone, with all its branches and all the muscle fibres supplied by them. It functions as a unit of motor activity. In all probability similar arrangements hold good for secretory and other types of neurone. Whatever the mode of termination, it is usual for the medullary sheath to disappear when it is reached. The neurolemma also sometimes disappears, leaving the terminal nerve fibre as a freely branching net of bare axis cylinder substance. In other instances the neurolemma may fuse with the wall of the cell in which the fibre terminates, as in skeletal muscle, in which the neurolemma and sarcolemma fuse together. The terminal nerve filaments thus come to lie directly in contact with, or may even penetrate into the cytoplasm of the cells at which they terminate, so that the relation between them is of the most intimate nature.

Nature of the Connection Between Neurones. The study of the development of the central nervous system in higher animals showed that from the embryonic cells, the neuroblasts, of the neural groove and crest there are formed by cell division the definite units in the construction of the nervous system which we now know as neurones. These retain their individuality from birth onwards and undergo no further divisions. It is now generally held, at all events for higher animals, that each neurone is a morphologically separate structure having no cytoplasmic continuity with other neurones, and being nutritionally distinct from them.

This conception of the neurone as a discrete structure is largely owing to the Golgi staining method, by which certain neurones are stained throughout all their branches, while adjacent neurones are quite untouched. Also, the facts of mammalian physiology seem mainly to be in favour of the neurone hypothesis. It is implicit in the present-day conception of the functioning of the nervous system that impulses can, under certain conditions, pass from one neurone to another, across some intervening junctional material. This junction is called the *synapse*. At the motor nerve endings in muscle there is also an apparent break of structure at the termination of the nerve in the muscle fibre, any continuity between nerve-ending and contractile substance being subserved by undifferentiated protoplasm. There is thus no difficulty in postulating a similar propagation across a synapse, from the axon of one neurone to the cell body or dendrites of another neurone.

If, however, the conception of the evolution of a nervous system from a continuous conducting protoplasmic network be correct, we should expect to find in the fully developed brain and spinal cord some traces of continuity throughout the whole system of neurones. The question as to the existence of anatomical continuity from neurone to neurone has been hotly discussed both for vertebrates and invertebrates. In some of the latter, evidence in favour of the continuity of neuro-fibrillæ from sensory surface to reacting tissue is strong. If we examine a mammalian nerve cell, such as a cell of Purkinje of the cerebellum, or of Clarke's column in the cord, we find that it is surrounded by a thick basketwork of fibres which are the arborisations and synaptic terminations of the axons which pass to the cell to enter into functional relationships with it (Fig. 118). It is probably not really a network, but a felt-work in which the fine fibres intertwine among each other without becoming actually continuous at any point. On the periphery of the cell itself another network has been described and is known as the Golgi network. According to Bethe the Golgi network on the one hand receives fibres from the encircling pericellular basketwork of axons, and on the other hand gives off towards the interior of the cell fine fibrils, which are continuous with the neurofibrillæ of the cell and pass out in its axon. There are physiological difficulties in the acceptance of this doctrine of continuity through the central nervous system. Thus, in the central nervous system, injury to an axon causes degeneration of the portion of the axon below the point of section, but the degeneration usually stops short at the synapse and does not spread into the next neurone. Exceptions to this rule occur, and the degeneration may affect several successive neurones—"trans-neuronal degeneration." A still more serious difficulty in the acceptance of this theory is the "Law of Forward Direction." It is possible that we may have to distinguish two types of nervous systems, viz.: (*a*) A neurofibrillar type, specially developed in invertebrata, with conduction in all directions; and (*b*) a synaptic type, which the Law of Forward Direction holds, of later evolution and forming the greater part of the nervous system of vertebrata.

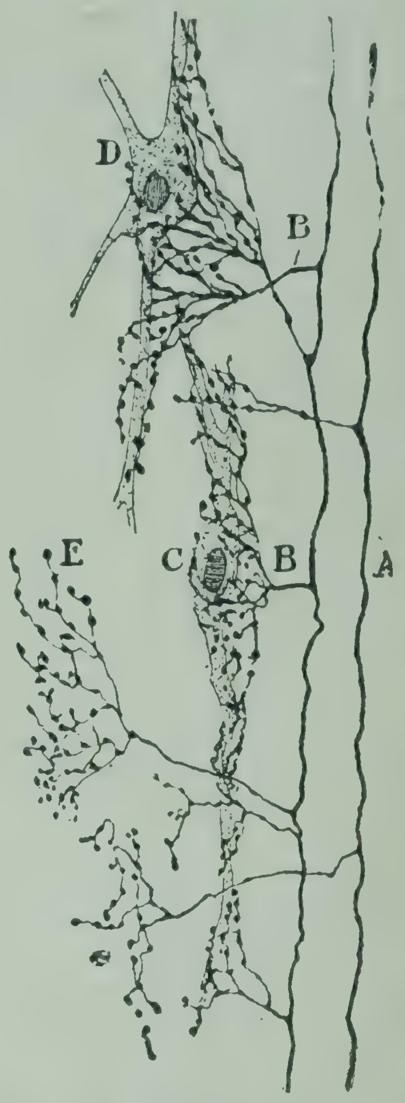


FIG. 118. Arborisation of Collaterals from the Posterior Root Fibres round the Cells of the Posterior Horn. (RAMÓN Y CAJAL.)

THE SYNAPSE. The synaptic arrangement by which communication can be established between contiguous neurones varies considerably in detail in different situations. In some cases the axon of one neurone and the dendrons of the next both break up into a finely interlaced mass of separate branches, comparable with the interlacing of fine branches of trees growing close together. More usually there is a plexus or basket of fine axon branches surrounding the body of the cell of the second neurone (Fig. 119) and terminating in small buttons or feet (*end-feet* or better, *synaptic knobs*) applied closely to the cell body or its dendrons (Fig. 120), or even to its axon-hillock. Although, in mammals, these synaptic knobs are generally minute (down to $0{\cdot}5\ \mu$), they may be very numerous (several thousand to a large soma), and cover an appreciable fraction (*e.g.* one-third) of the cell surface, to which they seem to adhere as if sticky. Typically, we may suppose, the knobs clustered upon a soma are derived from several—perhaps many—different axons, the fields of termination of the separate axons overlapping: but instances are known in which different axons terminate on separate areas, and sometimes by different shapes or sizes of terminal. Between the terminal knobs and the soma cytoplasm is a clearly stainable membrane, the synaptic membrane or synaptolemma, on the afferent side of which an aggregation of mitochondria is to be seen. Despite the variability of these anatomical arrangements there is no reason to suppose that the function of the

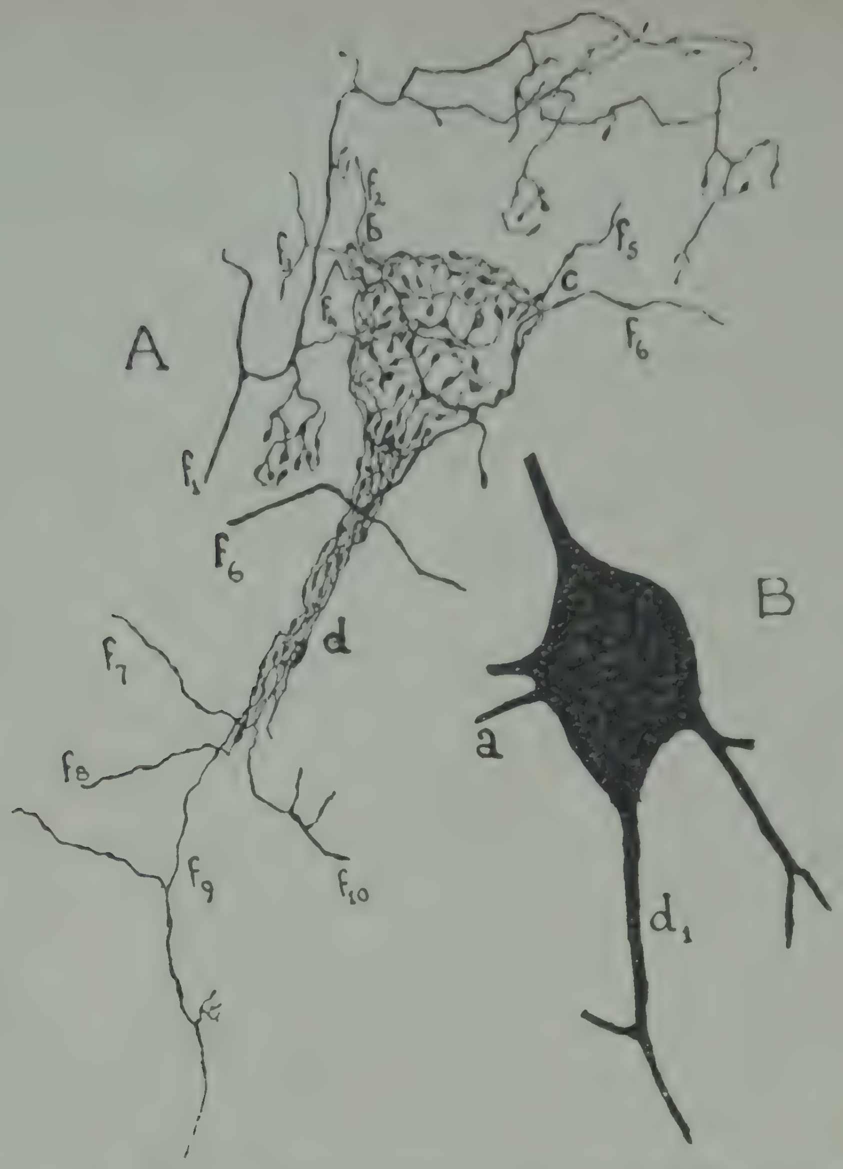


FIG. 119. A, synapses on a small internuncial neurone of the ventral ganglion of the acoustic nerve of a 20-day-old cat. B, a cell of the same type. *a*, axon; b.c.d,d_1, dendrites; f_1 to f_6, fibres having synaptic junctions with cell A. (LORENTE DE NÓ. *J. Neurophysiol.*, 1938, **1**, 195.)

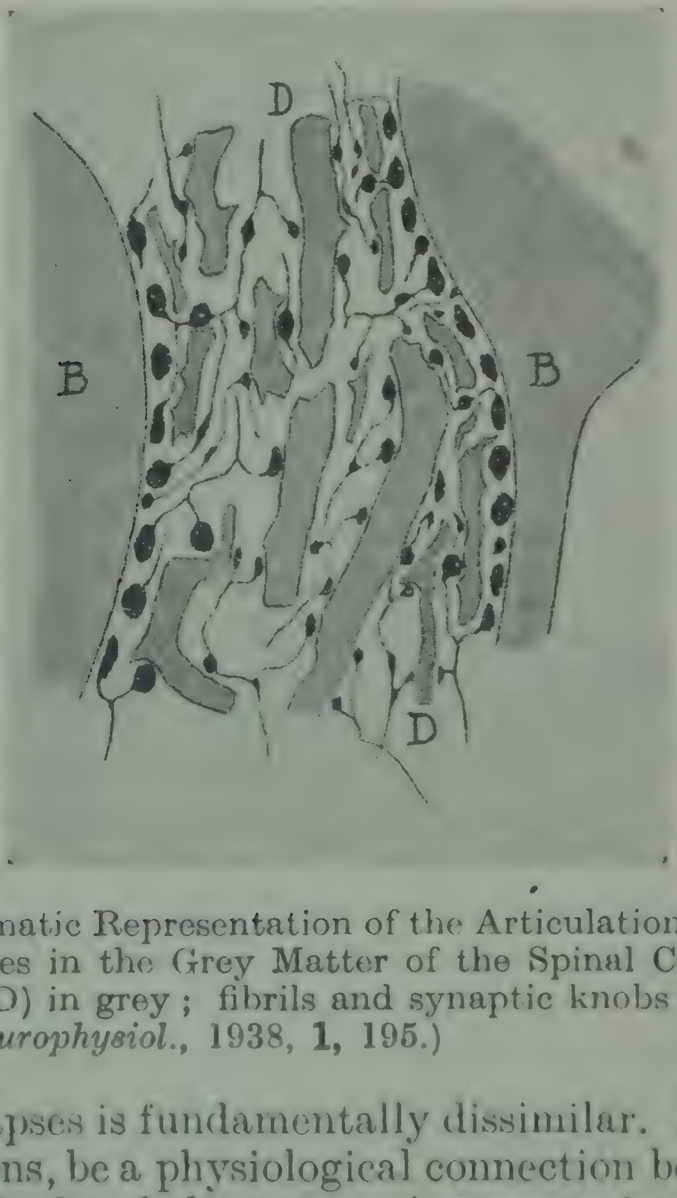


FIG. 120. Diagrammatic Representation of the Articulation of the Dendritic and Fibrillar Plexuses in the Grey Matter of the Spinal Cord. Cell bodies (B) and dendrites (D) in grey; fibrils and synaptic knobs in black. (LORENTE DE NÓ. *J. Neurophysiol.*, 1938, **1**, 195.)

different types of synapses is fundamentally dissimilar. In each case there can, under certain conditions, be a physiological connection between the axon of one neurone and the cell-body of the next. A nervous impulse can only pass in that direction, from axon to cell body, and so on to the axon of the second neurone. At the synapse there is usually a retardation of the nervous impulse called *synaptic delay*, which we may compare with the end-plate delay occurring on the passage of excitation from nerve to muscle Not infrequently a synapse is temporarily incapable of transmitting the excitatory state, a condition called *synaptic block*. But under no circumstances is the synapse capable of passing an impulse (or at least, does pass one, so far as we know) in the reverse direction from cell or dendron of one neurone to axon of the next, a rule which was called by Sherrington "*the law of forward direction*." In virtue of this the synapses act like valves allowing impulses to pass in one direction only. The probable reason for this is that the excitation time and intensity relations of soma and terminal are such that the soma surface is extremely unlikely ever to develop potential differences of such duration and intensity as to excite the knob around it.

It is an interesting subject of speculation as to the effects upon the cell of the arrival of impulses at synaptic knobs at different points in time and space, but as yet there is only a beginning of experiment in this field. Another interesting field for experiment and speculation is the enquiry into the fundamental nature of synaptic transmission.

When a stimulus is applied to any part of a neurone, *e.g.* at *x* in Fig. 121, A, it spreads throughout all its parts, and since the normal place for receipt of a stimulus is the synaptic area on the cell body, the normal direction taken by an impulse will be away from the soma and towards the axon, and along this centrifugally from its origin in the cell body to its terminal branchings *e.g.* from *a* to *b*, whence it will either reach an effector cell or another synapse, which latter, if patent, will be duly traversed, *e.g.* from *b* to *c*, so as to pass the excitation on to the next neurone.

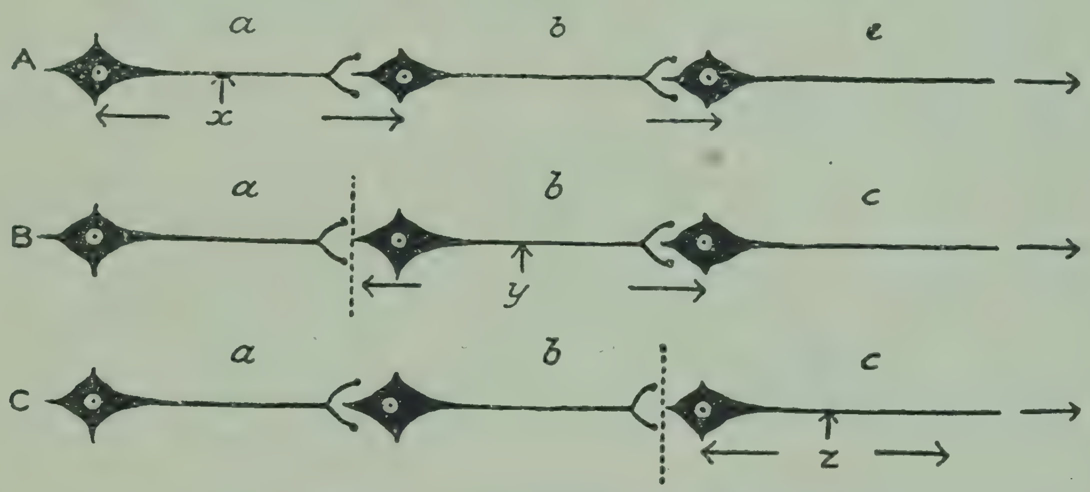


FIG. 121. Diagram to illustrate Forward Conduction.

If by any exceptional means an excitatory state be generated at an abnormal place, *e.g.* in the axon, as at *x*, *y* or *z* in Fig. 121, A, B, C, impulses will travel in both directions from the site of excitation, but only the forward or centrifugal one, *i.e.* which passes in the direction away from the cell body, will succeed in passing the synapse. The centripetal impulse will traverse the cell and dendrons but will be blocked " up-stream " at the synapse (Fig. 121, B, C).

Distributing, Commutator and Focussing Action of Neurones. From what has been said of the structure of the neurone and of the functions of the synapses, it is evident that the neurones play an essential part in distributing widely the impulses which they receive. This will be readily understood from Fig. 122, A, which shows that impulses received at a single neurone can be dispersed so as to reach many other neurones. This is called the *principle of divergence*. It is well illustrated in the optic nerve, each fibre of which sends branches to five or six cells of the lateral geniculate body.

A modification of this dispersing action is that seen in associative connection between neurones, as illustrated in Fig. 122, B. This is frequently of a reciprocative type, *e.g.* an excitation of neurone *a* may pass on an excitation through neurones *a′* and *b* to neurone *a″*, but the intercalated neurone *b* may now cause a block at the synapses leading to *c″*, with the result that as *a″*

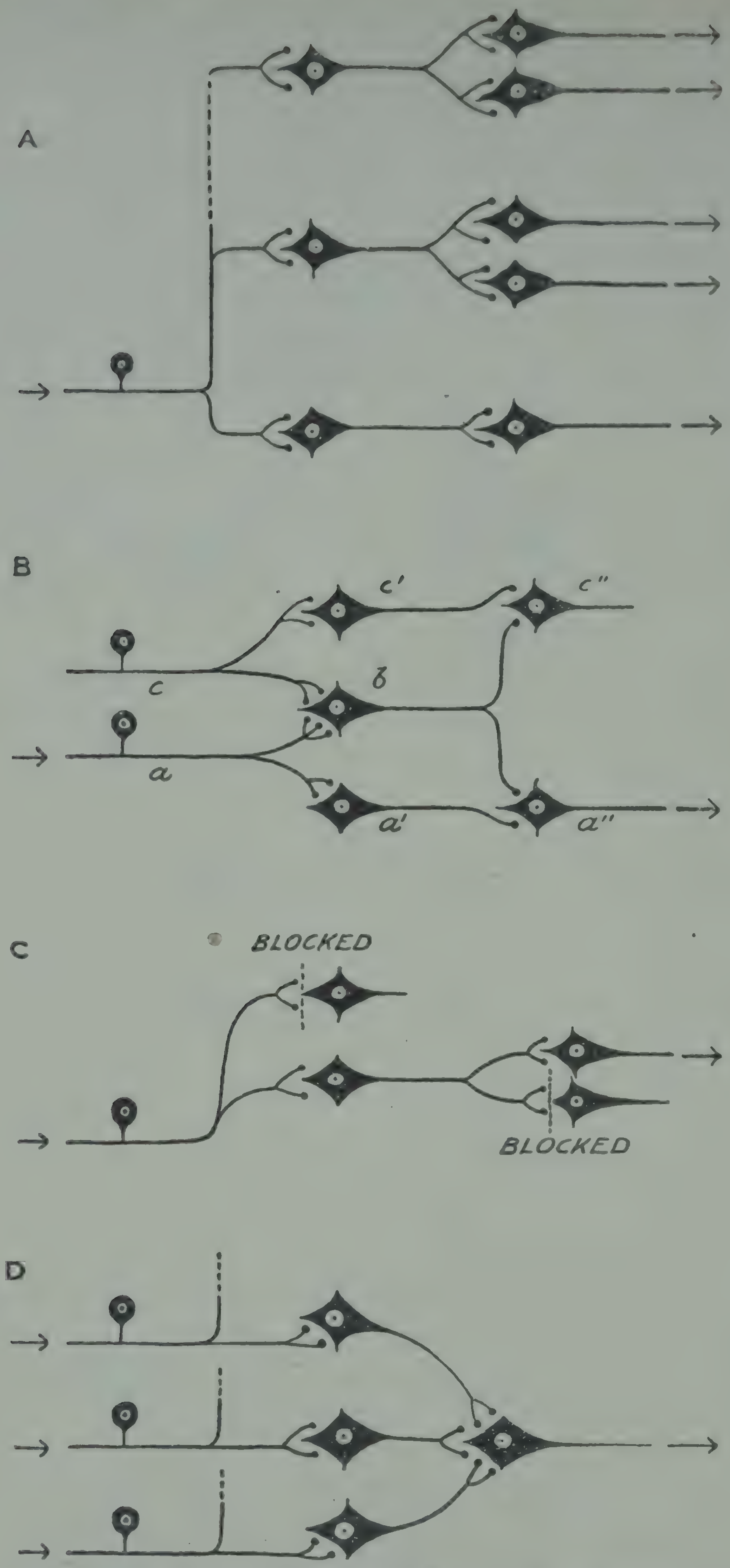


Fig. 122. Diagram to illustrate different connections of neurones.

enters into action, any activity in c'' is suppressed. The reverse would hold good for excitations applied to c, the synapse leading to a'' then being in a state of block owing to the action of neurone b.

A further function closely connected with the last is dependent on the fact that not all the synapses may be permeable to impulses, a result of which is that a number of alternative paths are open to impulses passing through a neurone, which thus acts like a commutator or railway junction (Fig. 122, C).

Lastly, owing to the multiple interconnections between neurones it is evident that a neurone may function in the opposite way to a distributor and may emit by its axon impulses collected or focussed from widely different sources (Fig. 122, D) to a "*final common path*." This is the *principle of convergence*. Instances of these three types of action are frequently met with in the central nervous system.

For Reference

Bodian (1942). Cytological Aspects of Synaptic Function. *Physiol. Rev.*, 22, 146.

THE MAMMALIAN NERVOUS SYSTEM

It is impossible to give here an account of the anatomy of the central nervous system, for which the reader should consult works on Anatomy and Neurology. All that can be attempted is to recall such outlines as are indispensable to our purpose.

Anatomically, the mammalian central nervous system is a modified tube consisting of the brain, which gives rise to the 12 pairs of cranial nerves, and the spinal cord, from which there emerge 31 pairs of spinal nerves. The hollow portions are modified to form the ventricles and the central canal of the cord.

The brain, or encephalon, is divided, largely on embryological grounds, into the rhombencephalon and the cerebrum, which are united by an isthmus. Of these the rhombencephalon comprises the myelencephalon, or medulla, and the metencephalon, or cerebellum and pons. The cerebrum includes the mid-brain or mesencephalon, the between-brain or diencephalon, and the cerebral hemispheres or telencephalon.

(1) **RHOMBENCEPHALON** (*IVth ventricle*)
  (*a*) *Myelencephalon* = Medulla.
  (*b*) *Metencephalon* = Cerebellum and Pons.
  Isthmus = Superior cerebellar peduncles and superior medullary velum.

(2) **CEREBRUM**
  (*a*) *Mesencephalon* = Mid-brain and cerebral peduncles (*Aqueduct of Sylvius*).
(*b*) *Diencephalon* = Hypothalamus, Thalamus, etc. (*Third ventricle*).
(*c*) *Telencephalon* = Basal ganglia (Corpus striatum), Amygdaloid nucleus and Claustrum Cerebral Hemispheres (*Lateral ventricles*).
Fore-brain.

The structures remaining after removal of cerebral cortex and cerebellar cortex are sometimes called the Brain Stem (Fig. 123).

The spinal cord, and the brain stem and cerebellum, *i.e.* the greater part of the brain apart from the cerebral cortex, are chiefly concerned with reflex functions of an instinctive type, and with conduction to and from the cerebral cortex. The cerebral cortex with its appendages is sometimes called the "new brain" (Neëncephalon), and is concerned with the highest functions of association, memory, etc. All its connections with the rest of the body and with the outer world are affected solely through the intermediation of the lower parts, the brain stem, cerebellum and spinal cord.

The details of the anatomical structure of the central nervous system are of little direct interest to the physiologist except in so far as they throw light on function. His attention is directed rather to the attempt to understand the different modes of neural activity, either as studies in themselves, or, more especially, in their effects as subserving the co-ordination of functions of the various organs of the body.

The functions performed by each of the parts of the central nervous system will be considered in the chapters which follow. Briefly, we may say that the function of the cord is to act both as a series of reflex centres and also as a conductor to and from higher parts of the nervous system. Histologically, it consists of a column of cells (the grey matter), which is H-shaped in cross-section, surrounded by bundles of medullated axons (the white matter). Of the nerve cells, some give rise to motor axons which leave the cord and go to muscles, some are connected with afferent nerves, to long sensory axons which convey impulses up the cord towards the cerebrum, or to structures below it, *e.g.* to the cerebellum, while some give origin to short-distance axons which correlate the action of neighbouring groups of muscles or subserve reflex acts.

The medulla contains the long path axons which are passing through it on their way to and from the cord, and also various accessory structures. In addition it contains important involuntary "centres" (control stations) for the heart, blood vessels, and for respiratory movements, as well as the nuclei (relay stations) of many of the cranial nerves. The mid-brain likewise lies on the main track; and also contains the relay stations of several cranial nerves. The ganglia at the base of the cerebrum act as sorting and relaying stations for impulses which are on their way to the cerebrum. They also have certain specific functions assigned to them which will be described later. The cerebellum lies off the main track and receives impulses both from the cord and medulla below and from the cerebrum above. It presides over the co-operation of muscles for the performance of skilled acts.

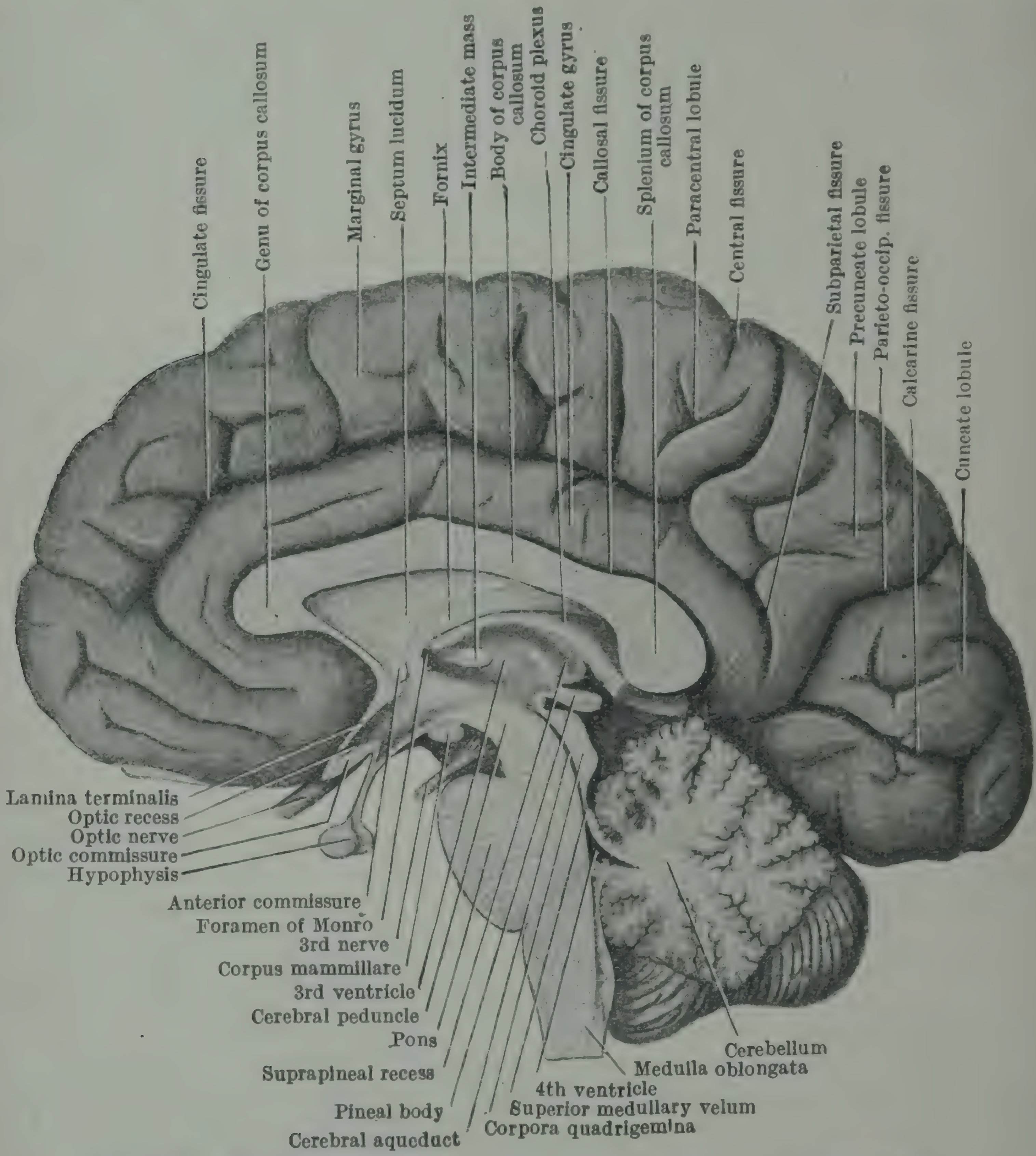

Cingulate fissure
Genu of corpus callosum
Marginal gyrus
Septum lucidum
Fornix
Intermediate mass
Body of corpus callosum
Choroid plexus
Cingulate gyrus
Callosal fissure
Splenium of corpus callosum
Paracentral lobule
Central fissure
Subparietal fissure
Precuneate lobule
Parieto-occip. fissure
Calcarine fissure
Cuneate lobule
Lamina terminalis
Optic recess
Optic nerve
Optic commissure
Hypophysis
Anterior commissure
Foramen of Monro
3rd nerve
Corpus mammillare
3rd ventricle
Cerebral peduncle
Pons
Suprapineal recess
Pineal body
Cerebral aqueduct
Corpora quadrigemina
Superior medullary velum
4th ventricle
Medulla oblongata
Cerebellum

FIG. 123. Medial Section of an Adult Human Brain. (J. SYMINGTON.)

The Brain Stem. It is usual in treating of the structure of the brain stem to consider it as a prolongation forwards of the spinal cord and as consisting, like this, of a central tube of grey matter surrounded by a tube of white matter. But at the fore end of the body have been developed the organs of special senses, which are the most important in determining the reactions of the animal in response to present or approaching changes in its environment. Hence we cannot expect to find in the brain stem the regularity of arrangement of grey and white matter which is met with in the cord, and the typical division of the grey matter into cornua becomes altogether lost. While some nerves take their origin from or terminate in the central tube of grey matter, in other cases the collections of nerve cells and fibres forming the nuclei of the cranial nerves have become more or less separated from it. Moreover, the central grey matter is by itself quite inadequate to deal with the flood of ingoing impressions entering the central nervous system through the organs of special senses, or to co-ordinate these with one another, or with those arriving from the skin and lower part of the body. In consequence, masses of grey matter, which have no representatives in the cord, make their appearance, and may be regarded as additional sorting or relay stations for the various kinds of impulses which determine the nervous activities of the animal.

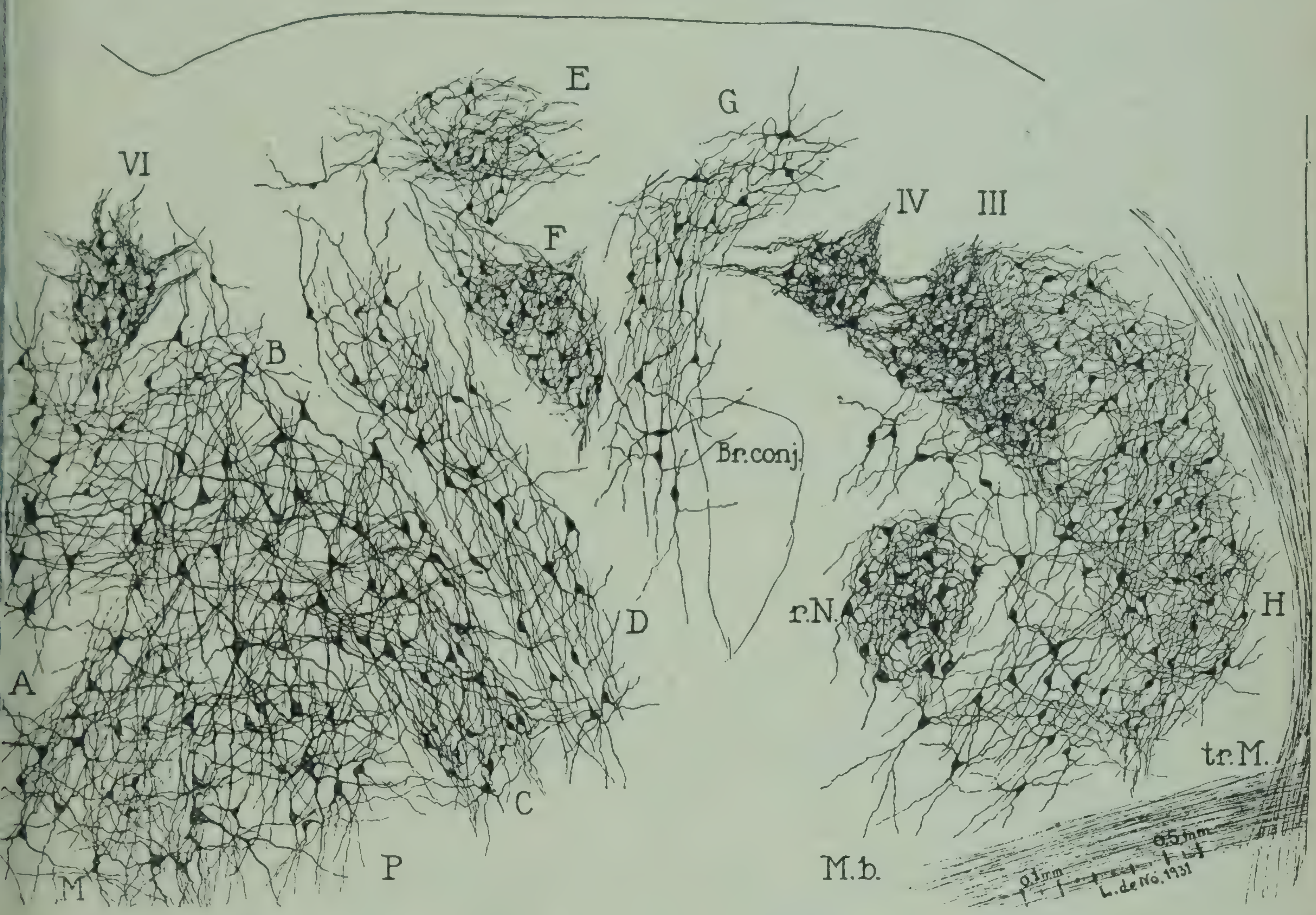


Fig. 124. Longitudinal Section through Medulla (M), pons (P), and Mid-brain (M.b.) of Mouse. III, IV and VI, the oculomotor nuclei; *r*.N, the red nucleus; Br.conj., crossing of superior cerebellar peduncles; A, B, C, D and H, various nuclei or neurone pools of the reticular substance, belonging to motor correlation system and having connections with the oculo-motor neurones. E, F and G are neurone pools belonging to the autonomic system; tr.M., tract of Meynert. (Lorente De Nó. *J. Neurophysiol.*, 1938, **1**, 207.)

The anatomical relationship between the parts of the central nervous system just mentioned may be gathered from a careful study of Fig. 123. The greater part of the diagram is occupied by the thick, much-convoluted cerebrum, the right and left halves of which are associated by the large mass of axons forming the corpus callosum. Below this will be seen a hollow funnel-shaped mass of tissue, which is narrowing as it passes downwards from cerebrum to mid-brain. In this mass the ganglia at the base of the cerebrum are situated. Past them the long path axons travel through the internal capsule, connecting cerebrum with mid-brain. The narrow neck of the mid-brain is well shown. Below this comes the pons which contains the long paths and which by its transverse fibres also associates the two halves of the cerebellum. Below the pons is the medulla, while the laminated cerebellum is seen lying behind.

Centres. Often the cells of the grey matter are arranged in groups of cells, of similar size and common functions, called *neurone pools*, and when such groups are sufficiently conspicuous, either anatomically by their size, or physiologically by their functions, they are usually spoken of as centres. Thus we can distinguish anatomically such groups as the oculo-motor nucleus, the red nucleus, etc., or physiologically a functional group such as the respiratory centre, micturition centre, etc. Those aggregations into physiological groups such as these latter are not necessarily anatomically definable, since they often consist of several closely co-ordinated groups of cells, which may be at considerable distances from one another: the unity is one of function rather than of position. Frequently, however, one group is of predominant importance and of precise location and is more clearly entitled to the name of centre, *e.g.* the cardio-inhibitory centre. Fig. 124 will give an idea of the actual anatomical relationship of different neurone pools.

The Autonomic Nervous System. In connection with the central nervous system is an important system called the autonomic nervous system. This, which subserves the control of most of the involuntary functions of a vegetative type, *i.e.* those concerned with the routine reflex adjustments of most of the principal systems of the body, is divisible into the sympathetic and parasympathetic systems. The *sympathetic system* is connected with efferent fibres from the thoracic region of the spinal cord, the *parasympathetic* with certain cranial and sacral nerves. The two systems have in common the possession of synaptic connections situated in pools called *ganglia* which lie *outside* the central nervous system.

The work of many of the medullary, and other " centres " which control important bodily functions is largely carried out through the mediation of the autonomic nervous system, which therefore is of much greater physiological significance than its modest anatomical build would lead us to expect.

REFLEX ACTION

The conception of the reflex arc was evolved early in the nineteenth century, though it is chiefly to Sherrington that we owe the initiation of researches which have made it a corner-stone of physiological neurology.

It is common knowledge that the application of many kinds of stimuli to different parts of the body produces an involuntary movement (Fig. 125).

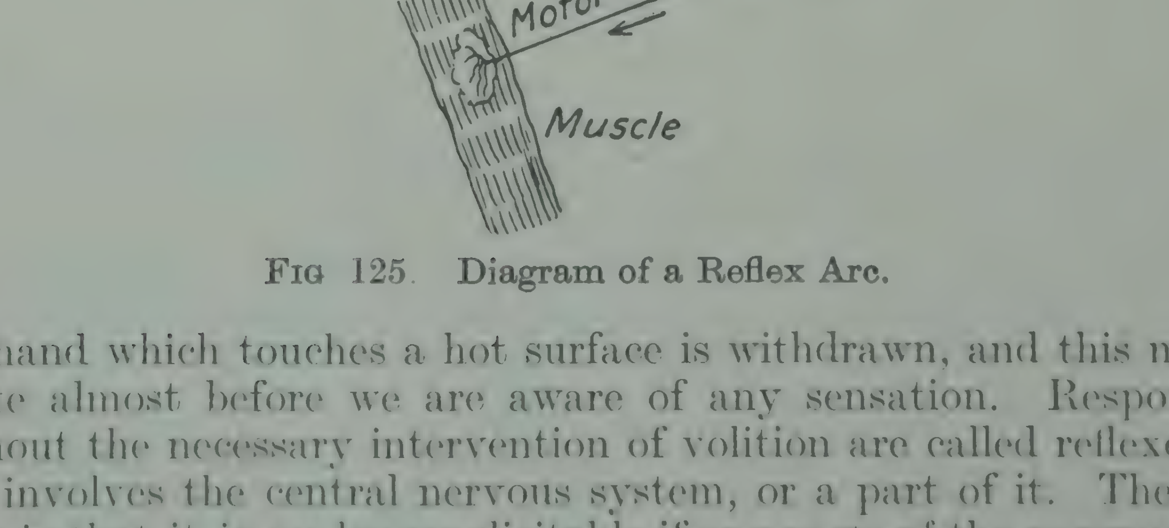


Fig 125. Diagram of a Reflex Arc.

Thus the hand which touches a hot surface is withdrawn, and this movement is complete almost before we are aware of any sensation. Responses that occur without the necessary intervention of volition are called reflexes if their execution involves the central nervous system, or a part of it. The criterion of a reflex is that it is no longer elicitable if any parts of the nervous path involved is destroyed. Sensation may accompany or follow a reflex action, but many important reflexes occur without the obtrusion of sensation into consciousness, as *e.g.* in the normal blinking of the eyelids at frequent intervals.

These automatic responses are of great variety, involving every system and organ of the body, and in a majority of cases they serve to protect, or otherwise further the interests of, the organism as a whole, or of the species. By their agency the various involuntary actions of the body under innumerable changes of conditions are properly effected and suitably co-ordinated, and its voluntary actions rendered more smooth and efficient. The secretions of glands and the movements of the viscera, the responses of the heart and blood vessels, and indeed all those bodily functions of which we are nearly always unaware, are largely controlled by delicately co-ordinated reflex actions.

Most of our detailed knowledge of reflex action, however, has been gained by the study of muscular responses, for the reason that they are the most objective and easily measured. As an example of a motor reflex in man we may take the knee-jerk. If one sits with, say, the right thigh crossed loosely over the left knee in such a way as slightly to stretch the quadriceps extensor, and then taps the right patellar tendon, a sharp extension of that leg at the knee-joint results. For this, as for all other reflexes, certain component structures are indispensable, and these components when reduced to their simplest terms of a one-fibre track form what is known as a reflex arc.

THE REFLEX ARC

A reflex arc comprises an afferent " limb," a centre, and an efferent " limb " (Fig. 126).

(1) The *afferent limb* of the arc consists of the receptors, and the afferent neurone. In the case of the knee-jerk, the receptors in the quadriceps muscle are stimulated by the tap. As a result of this they originate nervous impulses which enter the spinal cord by the dorsal roots of, say, the 4th lumbar nerve.

(2) In *the centre* (in our example in the 4th lumbar segment of the spinal cord) the impulses pass by a collateral fibre, *e.g.* from fibre B, direct to the ventral horn, and there, at a synapse (synapse 2 of Fig. 126) excites a large motoneurone. This is the simplest type, or *two-neurone arc*. In other less simple reflexes, *e.g.* some flexor reflexes, the collateral from a fibre such as A in Fig. 126, passes into the dorsal horn of grey matter, and so by synapse to a nerve cell, known as an *internuncial* cell (or interneurone, or relay cell) which is in turn excited ; the nervous impulse thus relayed passes by a short axon to the ventral horn and there, at a second synapse, excites a large multipolar motor cell. This is an example of a *multi-neurone arc*.

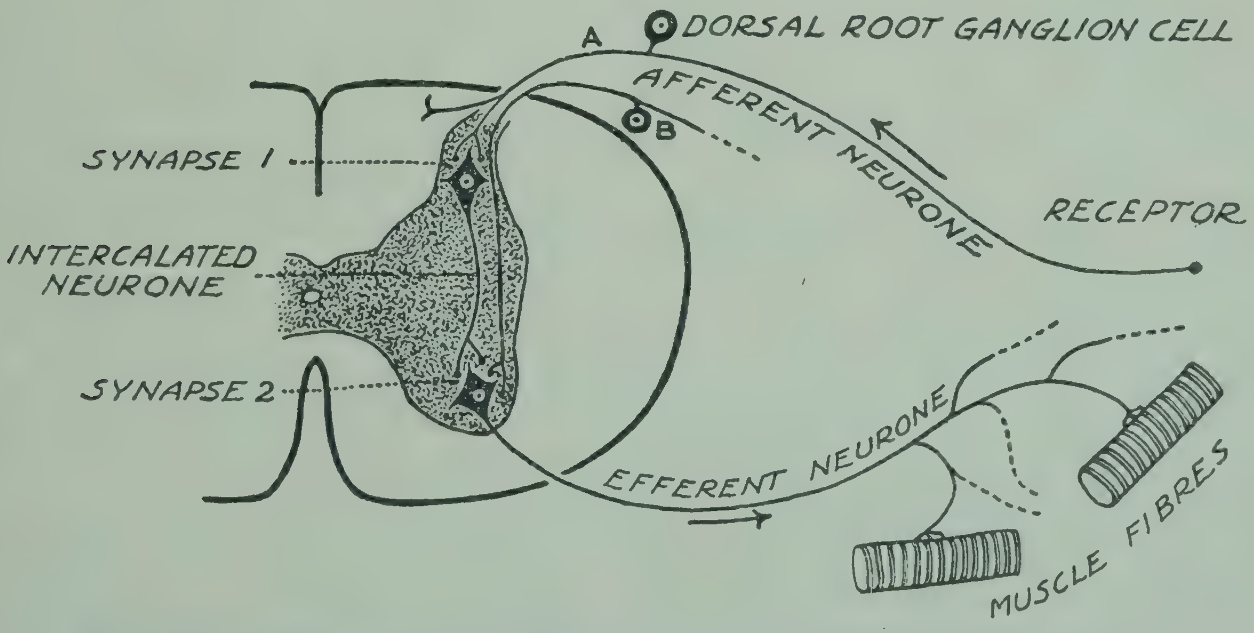


Fig. 126. Diagram illustrating the Reflex Arc in the Knee-jerk.

(3) *The Efferent Limb.* This consists of the motor nerve cell, its axon or motor nerve fibre, and the *effector organ* which effects the response (Fig. 126). The excited motor cell discharges a fresh relay of impulses along the motor

fibres which leave by the ventral root (4th lumbar) and so the excitation is ultimately conveyed to the muscle (quadriceps). The whole process for the knee-jerk occupies about 10 msec., most of which is taken up by the passage of the impulses along the nerve fibres.

We may substitute any other receptor organs, paths, centres and effector organs, but the general principle remains the same as that illustrated above. For example, we may consider the reflex secretion of saliva which follows when an acid is introduced into the mouth, as shown in Fig. 127. Here the afferent nerve is the VIIth cranial, the centre is in the medulla, the efferent path is by the chorda tympani (with another cell station peripherally) and the effector organ is a salivary gland. No doubt internuncial neurones are also placed in the arc.

Since the arc for a particular reflex only involves a limited amount of the central nervous system, it is possible, and often convenient, when studying reflexes in the laboratory, to remove parts of the central nervous system not involved. This is well illustrated in the "spinal frog" in which the brain has been destroyed, and in which spinal reflexes are easily studied. For example, if the frog is hung up by the head, and a toe of a hind leg is pinched, the leg is promptly drawn up in a co-ordinated flexor movement; if it is held so that it cannot move, the opposite limb moves vigorously instead. Again, if a scrap of filter paper soaked in dilute acetic acid is placed on one flank, the hind leg of the same side performs accurate movements to remove it; if that leg is held firmly, the opposite leg is brought into action.

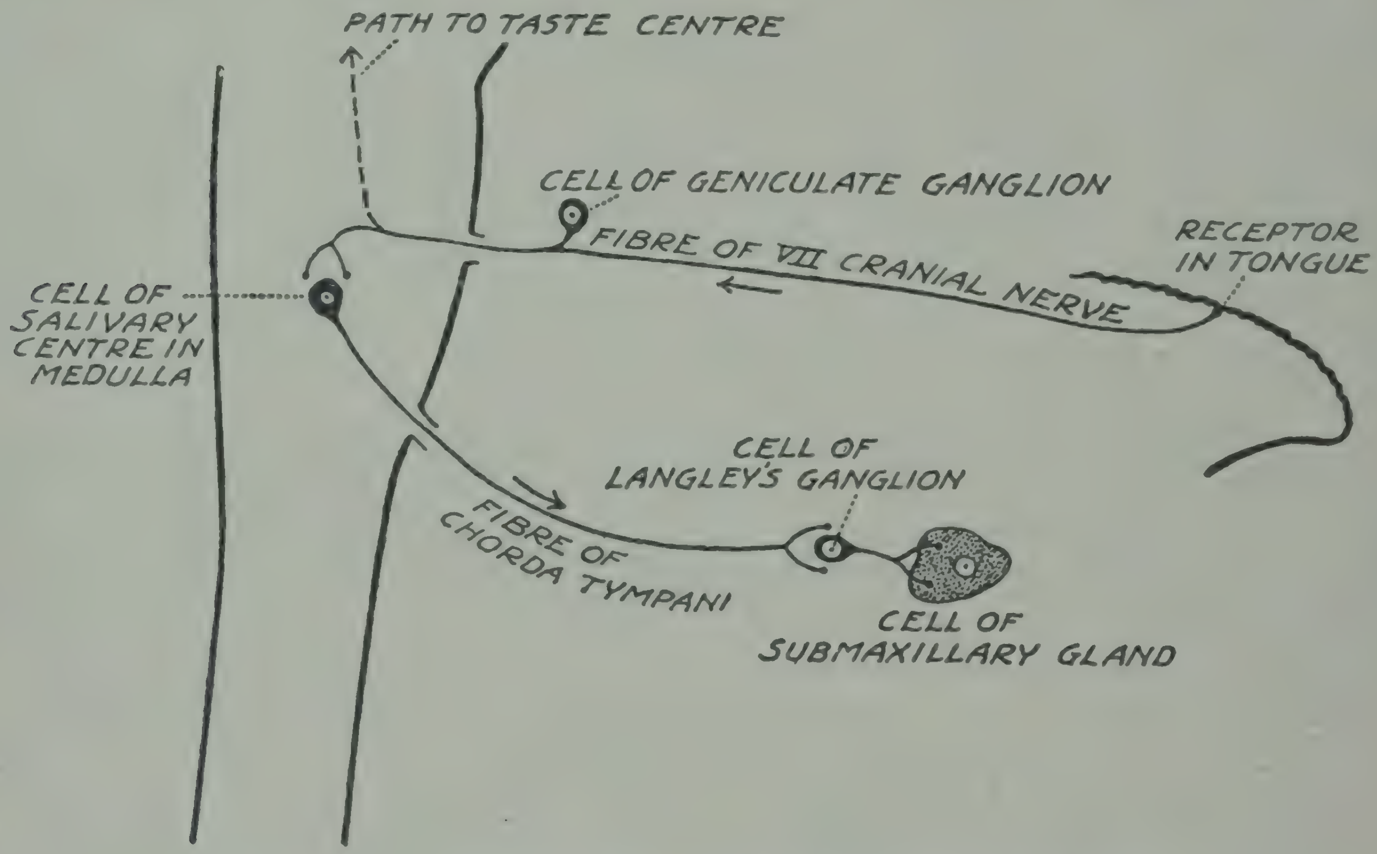

PATH TO TASTE CENTRE
CELL OF GENICULATE GANGLION
RECEPTOR IN TONGUE
FIBRE OF VII CRANIAL NERVE
CELL OF SALIVARY CENTRE IN MEDULLA
CELL OF LANGLEY'S GANGLION
FIBRE OF CHORDA TYMPANI
CELL OF SUBMAXILLARY GLAND

FIG. 127. Illustrating the Reflex Secretion of Saliva.

The spread, or *irradiation* of a reflex, as shown by this passage to the opposite side, illustrates the fact that alternative paths are available in the central nervous system, and a further spread is seen if both legs are restrained, when the fore legs enter into action.

Integrity of Reflex Arc. The necessity for the integrity of the reflex arc can also be illustrated with the spinal frog. If all the nerves of a limb are divided and the foot of that limb then stimulated, there is no response on either side, because of interruption of the afferent path. If the foot of the opposite side is stimulated that leg will be moved, but if its movement is prevented the denervated limb will not now be affected by irradiation because the efferent path is interrupted by division of its motor nerves. Lastly, if the spinal cord is destroyed with a wire, all reflexes are permanently lost, because the centres are absent. More elaborate experiments which have been performed on higher animals illustrate the same facts in more precise fashion. Thus if in a cat or dog all the dorsal roots or all the ventral roots are divided, no spinal reflexes can be elicited.

Integrity of Reflex Arcs in Man. In the human subject, the existence of certain reflexes supplies evidence for the integrity of the reflex arcs concerned, and is therefore of clinical importance. Thus the presence of knee-jerks shows that the 4th lumbar segment of the cord, with its afferent and efferent nerves, is intact, and in similar ways the following reflexes illustrate the integrity of the spinal segments indicated.

Segment.	Reflex.
C 5 and 6	Biceps-jerk
C 6	Supinator jerk
C 6, 7, 8	Triceps-jerk
Th 7–12	Abdominal reflex
L 1	Cremasteric reflex
L 2 to 4	Knee-jerk
L 4–5	Gluteal reflex
S 1 and 2	Ankle-jerk
S 2	Plantar reflex
S 4–5	Anal reflex

Other reflexes indicate the integrity of various higher parts of the central nervous system, *e.g.*, the contraction of the pupil of the eye to light shows that the IIIrd cranial nerve nucleus and the essential optic and efferent paths are intact.

In disease these reflex arcs may be interfered with in a variety of ways, some of which closely resemble the conditions introduced in laboratory experiments. Thus in tabes dorsalis, knee-jerks and other reflexes are lost because the dorsal spinal nerve roots suffer destruction, throwing the reflex arc out of action by loss of its afferent limb. In anterior poliomyelitis (infantile paralysis) the motor cells of the ventral horn degenerate, which causes paralysis of the affected segments by loss of the efferent limb of the arc. In syringomyelia the grey matter of the cord is generally affected, so that the whole central mechanism of the arc is lost.

CHAPTER XIV

THE RECEPTORS

THE receptors are the structures at which afferent nerve impulses are normally originated in response to appropriate, or adequate, stimuli. They are conveniently subdivided into exteroceptors and interoceptors. The *exteroceptors* convey impressions of the outer world, and are again divisible into telereceptors, chemical receptors and cutaneous receptors. The telereceptors are the end organs for vision and hearing, and will be specially considered under special senses, as also will the chemical receptors of taste and smell.

The *interoceptors* fall into the three groups of *proprioceptors*, *visceroceptors* and *chemical interoceptors*. The proprioceptors are end organs stimulated by the position or activity of the body itself, and visceroceptors similarly are situated in and give impulses arising from the viscera. Chemical interoceptors are present in the carotid sinus and certain other situations.

CUTANEOUS RECEPTORS

(For details, text-books of Histology should be consulted.)

The cutaneous receptors are of two types; viz. the free terminal ramifications and the specialised cutaneous end organs. Some of the principal types are shown in Fig. 128, but there are many others.

Free Ramifications. These are widespread and are usually non-myelinated at their terminations, or even throughout their course. Such terminal ramifications often form unmyelinated plexuses just beneath the epidermis,* and fine branches from this penetrate between or even into the epidermal cells. These free endings almost certainly subserve the pain sense and perhaps in some cases also touch.

Special Cutaneous Receptors. There are varieties of these, according to the various kinds or *modalities* of sensation, viz. touch, heat, cold and pain, *e.g.*,

Receptors for Touch. The chief varieties are: the hair follicle, the Meissner's corpuscle, and the Pacinian corpuscle (Fig. 128).

The hair follicles are supplied with a fine plexus of non-medullated terminal filaments, and these are stimulated whenever the hair is moved. *Meissner's corpuscles* are placed in certain of the papillæ of the skin. The fine non-medullated terminal fibres form a spiral imbedded in connective tissue. *Pacinian* corpuscles lie deep in the subcutaneous tissue of the hands and feet, but also in the joints and mesentery. The nerve fibre enters, loses its medullary sheath and becomes surrounded by an oval laminated structure of considerable size.

Receptors for Temperature Change. These are the Krause end-bulbs for cold (Fig. 128) and the Ruffini organs for warmth. In both of these there is a basket-like network of nerve fibres enclosed in an oval connective tissue capsule.

THE INTEROCEPTORS

Proprioceptors. The impulses from these receptors, often termed *kinæsthetic*, depend upon the movements or position of the body itself. They comprise the labyrinth of the internal ear, and the receptors of muscles, tendons and joints, and upon the impulses they furnish to the central nervous system the accurate pose and movement of the body largely depends.

---

* WOOLLARD. *Brain*, 1936, **58**, 352; *J. Anat.*, 1936, **71**, 54.

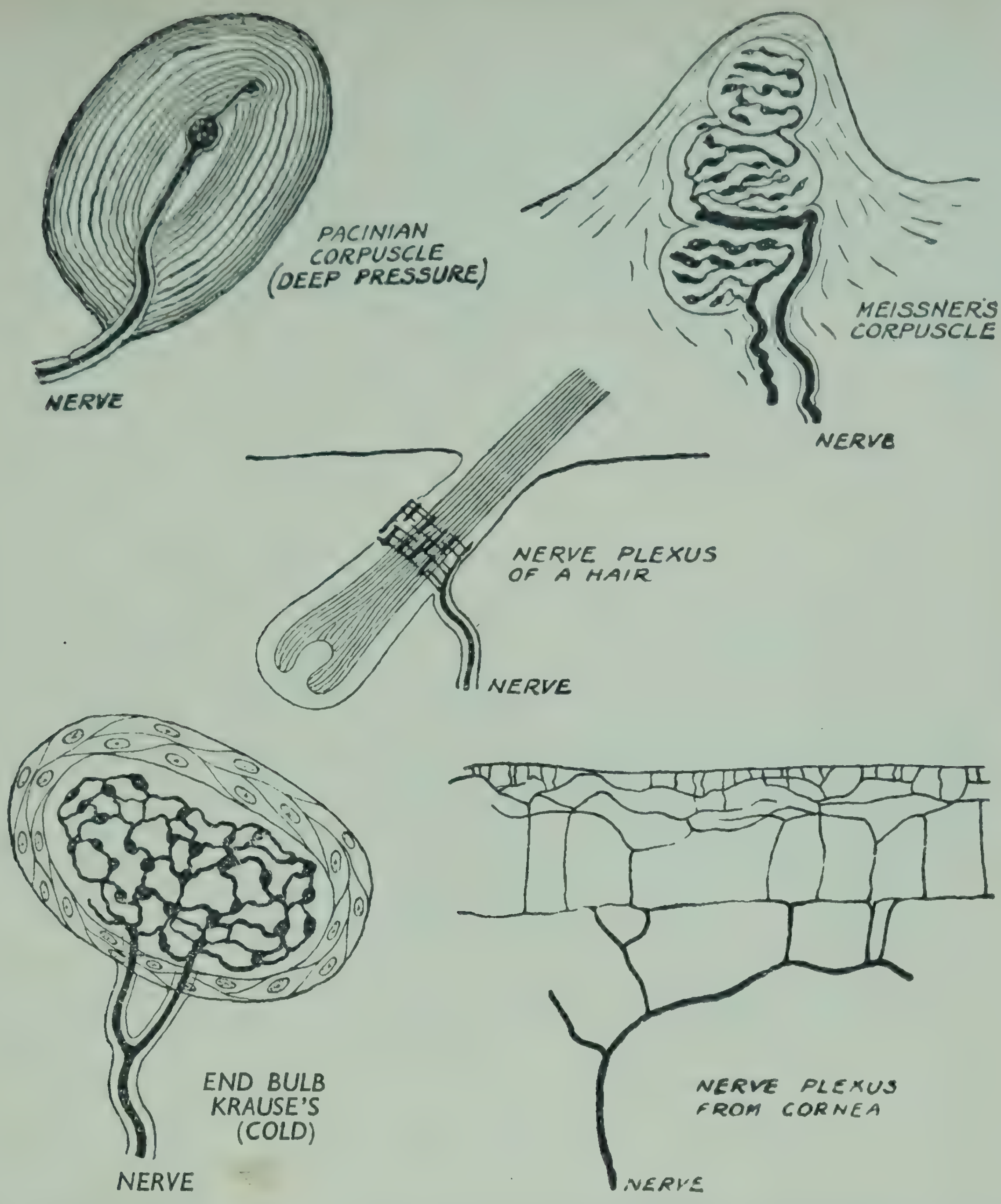

PACINIAN
CORPUSCLE
(DEEP PRESSURE)
NERVE
MEISSNER'S
CORPUSCLE
NERVE
NERVE PLEXUS
OF A HAIR
NERVE
END BULB
KRAUSE'S
(COLD)
NERVE
NERVE PLEXUS
FROM CORNEA
NERVE

Fig. 128. Skin End Organs. It has been supposed that corpuscles of the type of Meissner's corpuscle subserve touch sense, and that nerve plexuses, such as that illustrated for the cornea, are for pain.

The labyrinth consists of the *utricle* and *saccule* and the three *semicircular canals,* and will be considered in further detail in a later chapter. The receptors of the utricle and saccule are variously stimulated according to the *position of* the head in space, *i.e.* it is a static register, while the receptors of the three semicircular canals are differentially excited according to the *movement* of the head in the planes of space, and hence its impressions are kinetic in nature. The impulses aroused are conveyed from the labyrinth to the medulla by the vestibular division of the VIIIth cranial nerve.

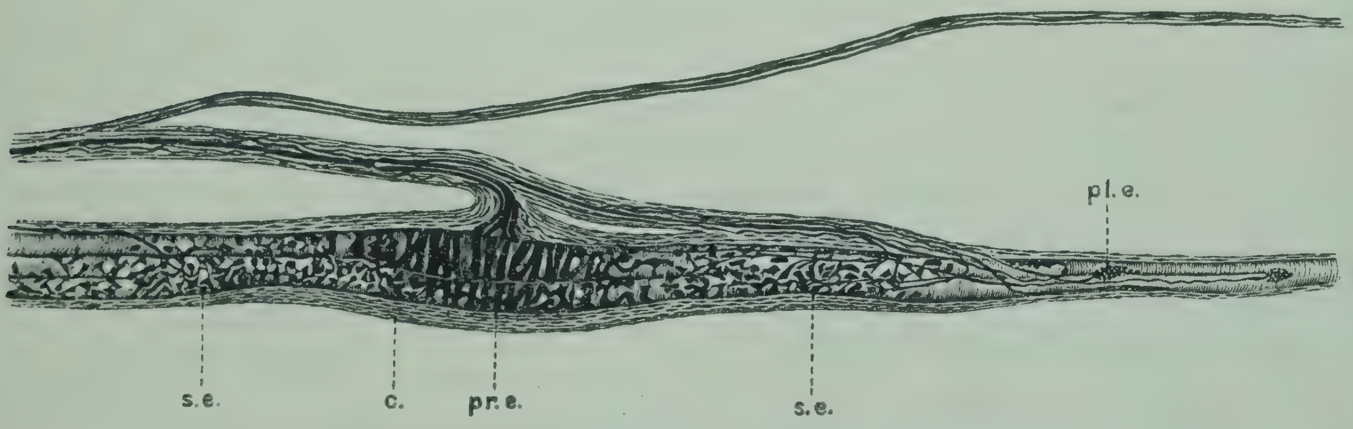

pl.e.
s.e.
c.
pr.e.
s.e.

Fig. 129. A Neuro-Muscular Spindle of the Cat. (Ruffini.)

c. Capsule.
pr.e. Spiral ending.
s.e. Dendritic ending.
pl.e. Plate ending.

Muscle Afferents. If the ventral roots supplying a so-called motor nerve are divided, and time allowed for degeneration to occur, it is found that as many as one-third of the fibres in the nerve remain undegenerated, *i.e.* are afferent fibres. The *muscle proprioceptors* are represented by the muscle spindles. In these structures a specialised type of muscle fibre, often split into several finer fibres, receives three nerve fibres. One is an afferent nerve fibre which, after penetrating the sarcolemma and losing its own neurolemma and medullary sheath, is spirally wrapped round the muscle fibres, the whole structure being enclosed within a capsule (Figs. 129 and 130).

Usually a second afferent fibre enters the capsule and forms dendritic branchings on the surface of the modified muscle fibres (Fig. 130). That these two types are both afferent is shown by the fact that their specialised endings disappear after removal of the dorsal spinal root-ganglia that supply the muscle in question. The third fibre is an efferent fibre with the usual motor end-plates. They degenerate on section of the ventral nerve roots concerned.

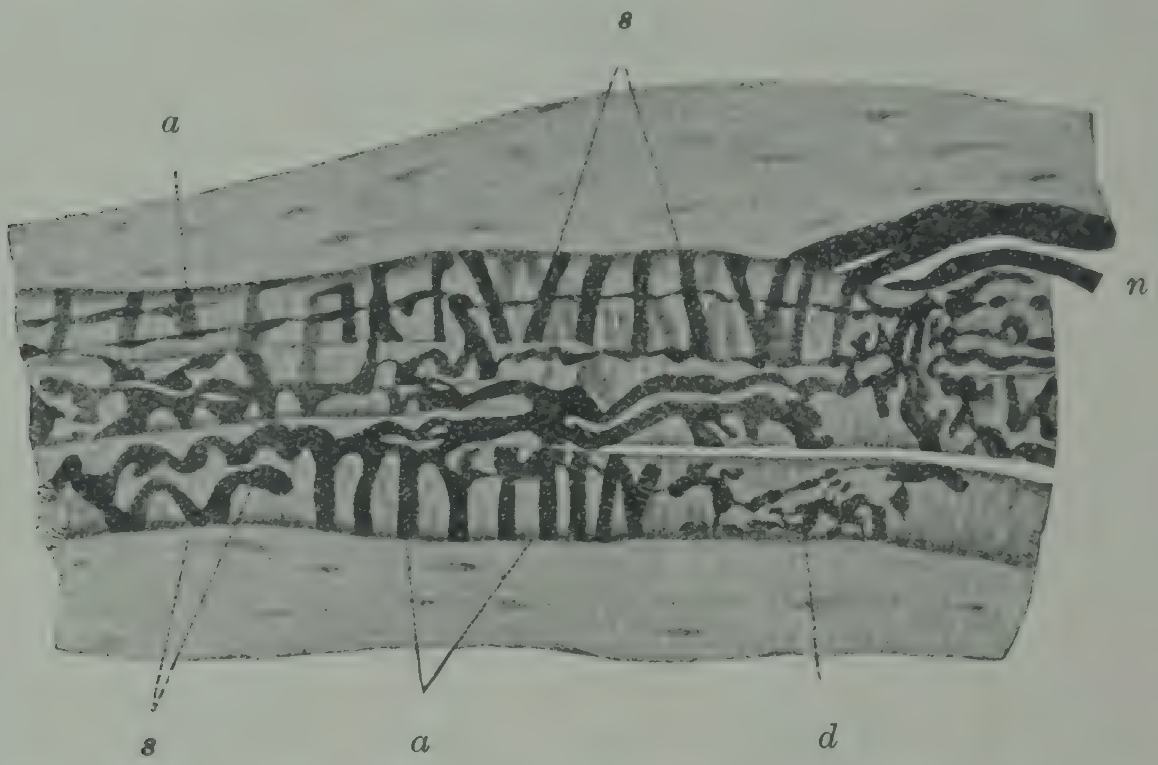


FIG. 130. Part of a Muscle Spindle more highly magnified. (RUFFINI.)

n. Nerve fibres passing to spindle.
a. Annular endings of axon.
s. Spiral endings.
d. Dendritic endings.

The muscle spindles are regarded as *stretch receptors*, *i.e.* they are excited when the muscle is elongated passively. When the muscle is contracting, any tension on the spindle fibres would be relieved by the support provided by surrounding fibres in a state of contraction.

The *tendon receptors* are represented by the Golgi tendon nerve-endings, which are found in the tendon near to its attachment to the muscle fibre. They are sensitive to tension, but since they are placed beyond the end of the muscle fibre, they will respond to a pull of any kind on the tendon, whether due to passive stretch or to actual contraction of the muscle. We therefore speak of them as *tension receptors.*

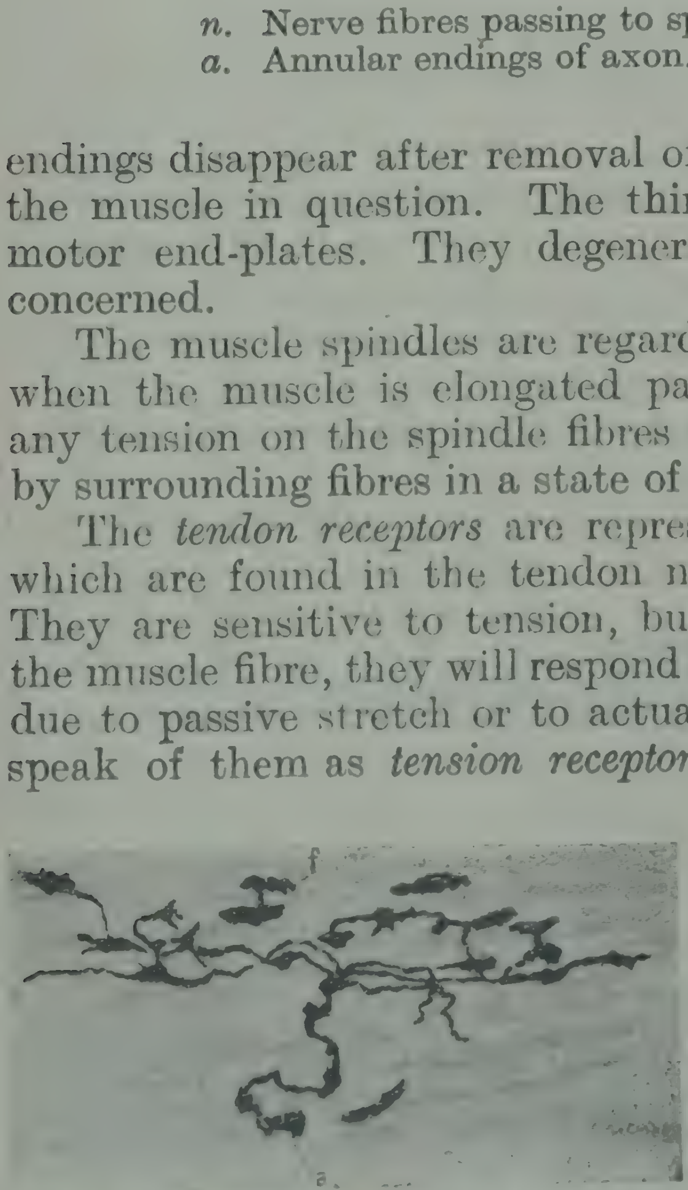


FIG. 131. Sensory endings in the adventitia of the carotid sinus in man: (*a*) arborising nerve fibres; (*f*) end organs. (C. Heymans, *after* de Castro.)

In addition to the above, receptors of other kinds will also be brought into play whenever movement occurs or posture is modified. Thus *Pacinian corpuscles* are found not only in the dermis but in tendon sheaths, joints, etc., and free nerve networks are also abundantly distributed to muscles, joints, tendons, etc.

Visceroceptors. The viscera and other internal parts of the body are not endowed with sensation in the ordinary acceptance of the term, but can be the seats of origin of a great variety of afferent impulses, including in some cases those capable of producing pain. Most of the hollow organs are provided with *stretch receptors* of some kind. These and other visceroceptors are often of the nature of dendritic branchings of non-medullated terminal nerve twigs. A remarkable instance of such a stretch receptor is in the carotid sinuses (Fig. 131), and aortic arch, from which afferent impulses arise when the pressure in the arteries is increased.

In addition to the stretch receptors, the carotid and aortic bodies in particular, are provided with *chemical receptors* which are affected by certain chemical changes in the blood passing through them.

Other important visceroceptors are the Pacinian corpuscles which are widely distributed over the peritoneal cavity.

THE EXCITATION OF RECEPTORS

Since it is the function of the receptors on suitable excitation to originate nervous impulses, we must enquire into the conditions and result of this stimulation. These differ in several important respects from those of nerve fibres. One difference lies in their specificity, *i.e.* they are so constructed as to respond only, or most easily, to one kind of stimulus, those of the retina to light, those of muscle to stretch, and so on.

Another difference is that the application of a stimulus to a receptor very readily leads to the origination of a *series or train of excitations*, which are duly passed on to the nerve fibre. This property is only feebly developed in nerve fibres themselves, so that stimuli applied directly to them do not, unless very strong, produce more than one nerve impulse per stimulus, even if this be a prolonged one. When a *prolonged steady stimulus* is applied to a receptor, however, a prolonged train of impulses is set up, and passed on to the nerve fibre, since each impulse in the receptor acts as a new stimulus to its nerve fibre.

Strength of Stimulus. Since the nerve impulse is an all-or-none phenomenon, the only effect of increasing the *strength of the stimulus* to a given receptor can be to produce a larger number of impulses along its nerve fibre. This results in the impulses following one another at shorter intervals, and also in the train continuing to be generated for a time after the stimulus has ceased. But another effect enters into the picture under most conditions in which a stimulus is applied to a tissue endowed with receptors, viz. as the stimulus is increased, an increasing number of receptors is excited. Hence an augmenting stimulus normally leads to an increasing rate of discharge along an increasing number of nerve fibres.

On account of this fact, the detection of the impulses initiated by a given stimulus requires, in practice, a special technique, which we owe to Adrian and Zotterman,* and to Adrian and Bronk,† by which the nerve impulses traversing a single nerve fibre are recorded. Adrian and Zotterman achieved this by isolating one receptor (spindle of sterno-cutaneous muscle of the frog) by cutting away the others; Bronk ‡ by dividing all except one of the afferent nerve fibres that led away from the sensory tissue (muscle or skin); Matthews § by using a tiny muscle (of a frog's toe), which had only one spindle.

In each case the spike potentials along the single nerve fibre were recorded, after amplification, and the complication due to the recording of impulses along many nerve fibres was avoided. By similar methods it has been possible to obtain records of the nervous impulses (spike potentials) which pass along many kinds of afferent fibres on appropriate stimulation of their receptors, and in all cases the results are essentially similar (Fig. 132).

---

* ADRIAN and ZOTTERMAN. *J. Physiol.*, 1926, **61**, 151.
† ADRIAN and BRONK. *J. Physiol.*, 1928, **66**, 81.
‡ BRONK. *J. Physiol.*, 1929, **67**, 17.
§ MATTHEWS. *J. Physiol.*, 1931, **71**, 64.

When the strength of the stimulus applied to the receptors is raised, the frequency at which they respond is increased up to a certain maximum, which

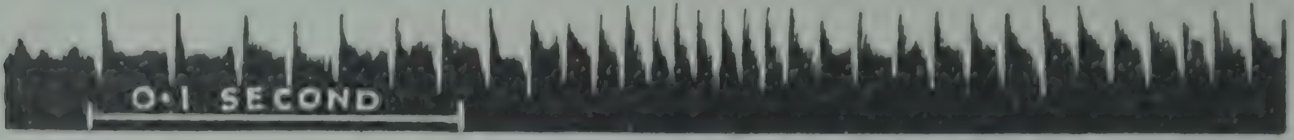


Fig. 132. Action potentials (*i.e.* nerve impulses) in a single nerve fibre from a frog's muscle spindle, caused by stretching muscle. The waves are all similar in size, but increase in frequency as the tension increases and then decline slowly when the tension is constant. (Tsai.)

in some receptors may reach as much as 300 per second ; this is about the limit at which a nerve fibre can fully respond to direct stimuli, since the absolute and relative refractory periods of the fibre are about 3·5 msec. It follows that the refractory periods of the receptors themselves must at most be of this order.

Adaptation. All nervous structures exhibit the phenomenon of adaptation, *i.e.* the responses to a stimulus become less and less frequent with time. Nerve fibres have this property so strongly developed that each application of a stimulus usually only produces a single nerve impulse. With receptors more than one excitation usually results from each stimulus, and they are capable of responding to a continuous stimulation by firing off a long train of responses which pass on to the nerve fibre. Thus the stretch receptors of muscle, or the touch receptors of the skin, when subjected to appropriate mechanical stimulation fire off a series of nerve impulses into their respective nerve fibres. The receptors, however, do show adaptation, as a result of which the train of impulses, at first in quick succession, soon shows a slowing down in frequency. In this respect the different receptors show considerable differences. In muscle receptors the rate at which impulses are generated falls off at first and then reaches a fairly steady rate of discharge (Fig. 133) ; whereas skin receptors adapt much more rapidly and may in spite of continued stimulation, ultimately only discharge very slowly. This corresponds to our own subjective experience. When the skin is lightly touched, as, *e.g.* by clothing, the sensation is much

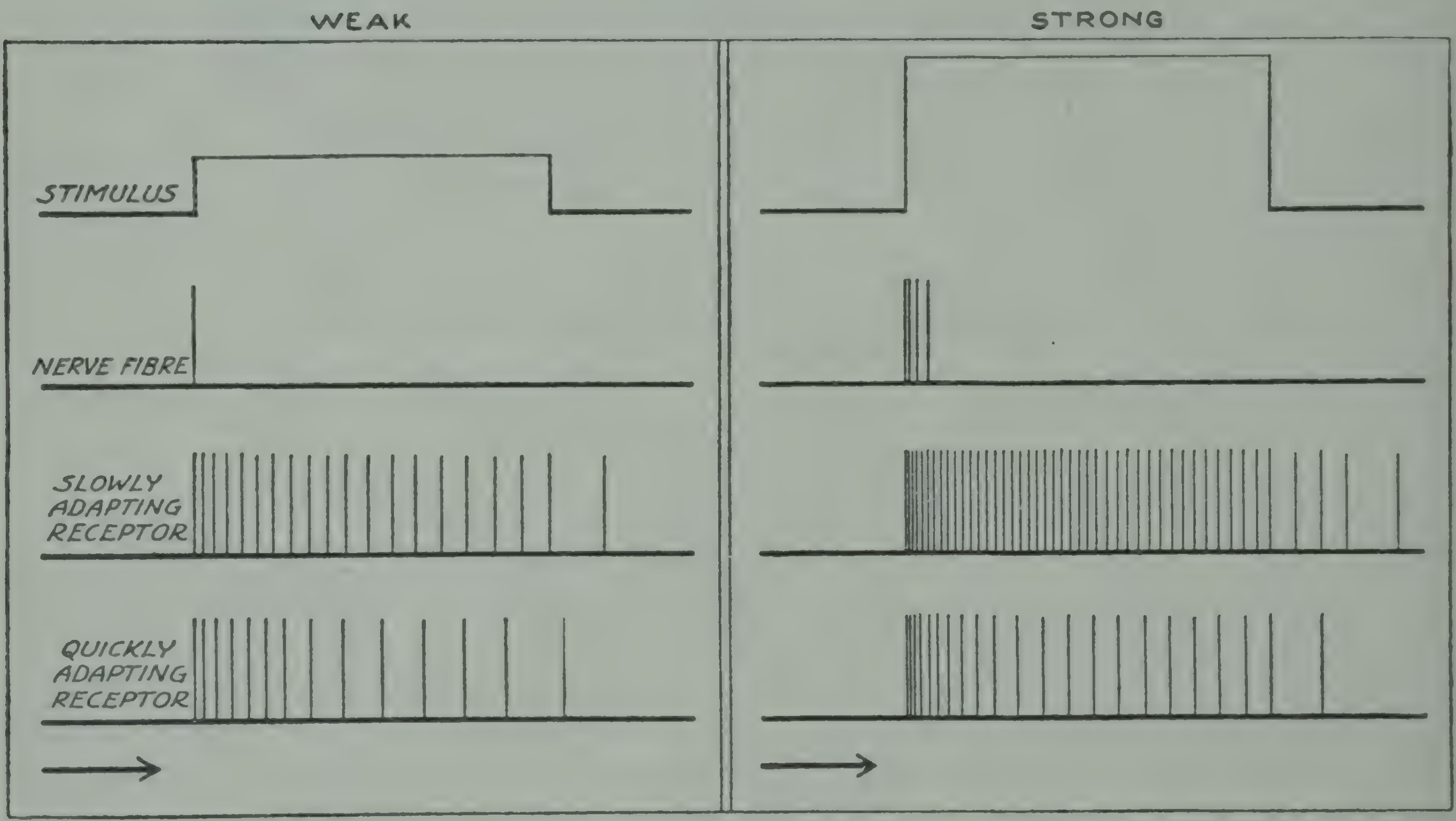


Fig. 133. Diagram illustrating the effects of continuous Stimulation, Weak and Strong, on Nerve Fibre and on Receptors. The vertical lines represent nerve impulses. Nerve fibres give only one impulse for weak stimuli or two or three for very strong ones. Receptors discharge trains.

more intense at first than it subsequently becomes—otherwise the wearing of clothes would be intolerable.

NATURE OF ACTION OF RECEPTORS. Why and how the various receptors respond to their appropriate stimuli, and to those alone, is an interesting question, but one to which no definite answer can at present be given. Whatever the manner in which the state of excitation is brought about, the ultimate result is the same, viz. the generation of trains of nerve impulses along the fibres whose endings they form. It is not improbable, therefore, that the effect on any end-organ of its appropriate stimulus is in all cases the same physico-chemical process. Observations on the retinal terminations suggest that this change is a surface *depolarisation*,* and this view would be acceptable, since it brings the process of excitation of nerve-receptors into line with that of other structures, and particularly with that of nerve cells, which, like the receptors, also generate impulses. Two facts are in accord with this view: artificial depolarisation by an electric current causes excitation; and an impulse sent backwards (antidromically) into the receptor by artificial stimulation of its nerve fibre causes the receptor to enter for a short time into a refractory state,† and so interrupts its normal rhythm of discharge.

---

* Hartline. *J. Cell. Comp. Physiol.*, 1932, **1**, 277.

† Matthews. *J. Physiol.*, 1933, **78**, 46.

CHAPTER XV

THE SPINAL CORD

AND ITS RELATIONSHIP TO THE HIGHER PARTS OF THE CENTRAL NERVOUS SYSTEM

It is not strictly possible to consider the central nervous system as a series of separate parts, the functions of which when added together make up the functions of the whole. Rather we should aim at understanding how the entire system functions. Keeping this end always in view, however, much that is of value can be learned by considering the separate functions of the parts. We shall therefore in the first place study the functions of the spinal cord in brief fashion, because it is simpler in structure, and because on that account much can be learned about neural physiology that can subsequently be applied to the understanding of the functions of the higher parts.

In man the spinal cord gives off, symmetrically on either side, thirty-one pairs of nerves. Each nerve has two roots, an anterior, or ventral, root and a posterior, or dorsal, root. The ventral roots are efferent (*e.g.* motor) in function, and the dorsal ones afferent (sensory). Both types of root contain fibres of widely differing diameters. A mixed nerve consists of fibres which are connected to both these roots. The anterior root fibres emerge from the cord in a series of rootlets; the posterior root fibres, on their way into the cord, pass first to the posterior root ganglion and then enter the cord as a compact bundle. On section the cord is seen to consist of an H-shaped core of grey matter surrounded on all sides by white matter. The former contains the nerve cells, the latter is made up of nerve fibres which are devoid of neurolemma and are connected into bundles held together by neuroglia.

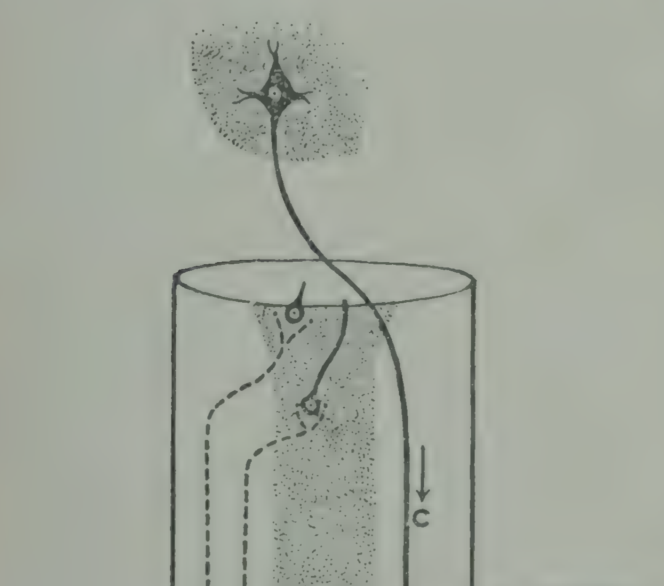


Fig. 134. Diagram to illustrate Method of tracing Tracts by Degeneration. A, fibre of an exogenous ascending tract; B, fibre of an endogenous ascending tract; C, a descending fibre. Degenerated tracts shown by broken lines.

THE NERVE CELLS OF THE CORD

The nerve cells of the grey matter of the spinal cord are all multipolar cells; they are not uniform in size or distribution, but tend to be aggregated in pools or columns, the cells of each column having approximate similarity of function. The cells are of Golgi types I. and II. The type II., with short axons, are found only in the dorsal horn, and chiefly in its lateral basal group. Their short axons pass towards the ventral horn of the same or the opposite side. The type I. cells usually send out their axons into the tracts of the white matter. Important groups of the posterior horn are the cells of the *substantia gelatinosa Rolandi*, and at the base of the posterior horn is the *Clarke-Stilling column* (or mesial basal group). The cells of the substantia gelatinosa Rolandi send out axons mainly into the spinothalamic tracts of the opposite side. The axons of the cells of the Clarke-Stilling column pass out into the dorsal and ventral spino-cerebellar tracts and so reach the cerebellum, mainly of the same side (Fig. 137).

In the lateral portions of grey matter, in the thoracic and sacral regions, are cells forming the *lateral splanchnic column* which give origin to axons that pass out *viâ* the ventral roots and white rami communicantes, to the sympathetic chain (Fig. 137).

In the ventral horn of grey matter are three or four groups of cells forming the *motor cell columns.* Their axons pass out into the ventral spinal nerve roots, and so to the spinal nerves.

THE WHITE MATTER

The white matter of the spinal cord represents the conducting matter of the cord, and contains fibres of four types: (i.) long path afferent axons, such as those of Goll and Burdach in the posterior columns, forming the main tracts to higher parts, (ii.) long path efferent axons, leading down from higher parts, *e.g.* the pyramidal tracts, (iii.) association or intrinsic axons, which connect neighbouring segments of the cord, (iv.) commissural fibres which connect opposite halves of the same segments of the cord. The fibres are of varying size, and some are non-myelinated.

Much of the white matter is divisible into *tracts,* consisting of long fibres of similar function and destination, and these again are distinguished as *afferent and efferent tracts.* The tracts are mapped out mainly by the degeneration method. If a tract is divided, the parts of the axons cut off from their cells will degenerate, *i.e.* right up to the next synapse, and after some days may be distinguished by appropriate staining methods. The same thing happens if the whole spinal cord, or other part of the central nervous system, is cut across; the axons from cells below the lesion (ascending tracts) will be degenerated in the part above the section, and those with cells of origin above the lesion (descending tracts) will be degenerated only below the section (Fig. 134). If the dorsal nerve root be divided proximal to the ganglion, *i.e.* between the ganglion and the cord, it will be clear that the whole of the course of the axon within the cord will degenerate, and this method provides a way of differentiating those afferent tracts which are directly derived from the dorsal root fibres (A in Fig. 134) from those of secondary origin in the grey matter of the cord (B in Fig. 134).

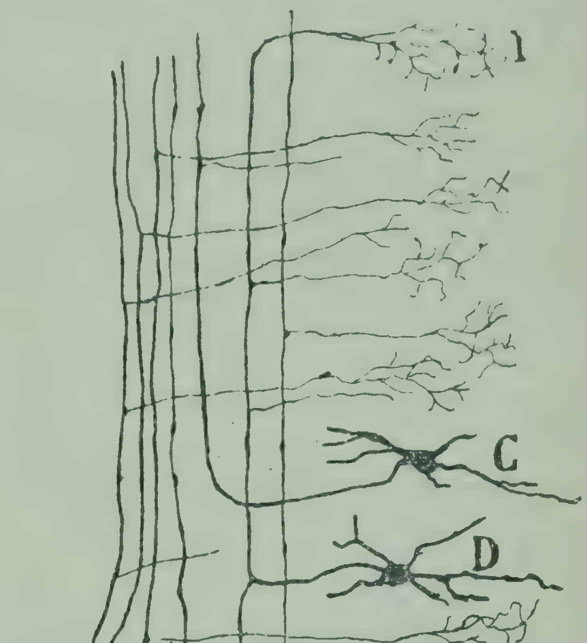


Fig. 135. Longitudinal Section of Spinal Cord of Chick, showing bifurcation of Dorsal Root Fibres and the passage of their Collaterals into the Grey Matter. Three cells of the Dorsal Horn are also seen sending their axons into the dorsal columns of white matter (short intraspinal neurones). (Cajal.)

Degenerating myelinated fibres are picked out by the Marchi method. In studying fibre degenerations in the C.N.S. the non-myelinated fibres must also be taken into account: these are stained by a silver method if undegenerated.

In some instances the divided fibres still connected to the cell also degenerate (retrograde degeneration); usually the cells of origin of divided tracts undergo chromatolysis. Hence the pool of cells from which a tract arises (*e.g.*, tract B in Fig. 134) can often be identified. Cell and fibre counts will indicate the extent to which a given neurone pool contributes to a given fibre tract.

Another method of following the course of impulses in the central nervous system is by the use of microelectrodes and amplifiers. With sufficiently small electrodes the passage of impulses can be followed along separate fibres, using the natural stimuli applied to skin, muscle, etc., and the velocity of the impulses, which is related to the fibre diameter, can be estimated. Stimuli can also be applied direct to selected tracts in the C.N.S., and the resultant impulses traced. This method, sometimes called *electroneuronography*, has opened up new fields of study.

In some parts of the central nervous system the connection between one set of neurones and the next can be determined by what is called *transneuronal degeneration*; *i.e.* when one neurone is removed the next one degenerates. Thus if a part of the retina is excised, the cells which connect with it in the lateral geniculate body degenerate.

The chief paths of physiological importance are briefly as follows :—

I. Afferent Paths. Each dorsal root fibre, on entering the cord, gives off a short descending branch and an ascending branch (Fig. 135), each of which has numerous small collaterals which enter the grey matter of the dorsal horn and there form synaptic connections with cells in one of the pools. The ascending branch may, (*a*) after a short course, similarly terminate in the grey matter, by synapse with a cell from which an axon arises, to pass out again into the white matter as a fibre of a secondary tract; or (*b*) pass directly as an "exogenous" tract into the white matter of the posterior columns (Fig. 137). On entering the spinal cord the fibres become grouped into tracts according to their origin, destination and functions.

As regards origin, the fibres are arranged within their tracts in an order corresponding to the dermatomal sequence. As regards destination and function, fibres of like function and destination travel in the same tract: thus, broadly, the posterior columns convey fibres dealing with proprioception and fine conscious discriminations, while the anterolateral columns are concerned with pain, temperature and sex sensations, and with unconscious afferents.

It has been argued by Stopford * that the afferent nerve fibres fall naturally into two groups. One group is connected with defensive functions, is phylogenetically the older, gives sensory responses of an all-or-nothing character as regards their affective qualities, and has important central connections with the thalamus. This group of fibres (spino-thalamic) crosses in its path in the cord. It includes sensations of pain, of extremes of temperature, and of pressure. The other group (posterior columns), less old, uncrossed in the cord, is essentially cortical in its terminations, and is concerned with finer discriminations of localisation, discrimination of position and passive movements, fine differences of temperature, and appreciation of lightest forms of touch.

Fig. 136 shows the main tracts of the spinal cord and Fig. 137 illustrates the mode of formation of the chief afferent paths, which is as follows :—

(1) FINE SKIN SENSATIONS, *e.g.* light touch, light or heavy pressure, fine localisation and discrimination, stereognosis, pass up at first in the *dorsal columns*, and continue there to reach the nuclei gracilis and cuneatus. The arrangement of the fibres is according to the order of dermatomes, the sacral fibres nearest the mid-line, the cervical ones furthest from it. Some of them, however, of cruder function, end synaptically at cells of the posterior horn, from which axons pass out to the *ventral spino-thalamic* tract of the opposite side. [This joins the spinal lemniscus and ends in the thalamus.]

(2) AFFERENT IMPULSES RELATED TO MOVEMENTS. These may be conscious, *i.e.* give rise to sensations, or unconscious; most of the latter are related to the maintenance of posture.

(*a*) The impulses subserving *conscious sensations of movement* enter at the posterior root, ascend in the columns of *Goll and Burdach* until they reach the *nuclei gracilis* and *cuneatus* in the medulla. [The cells in these nuclei give off second order relay fibres which, crossing in the sensory decussation, ascend in the mesial lemniscus (fillet) to the thalamus. Third order fibres originating in this convey the impulses to the cerebral cortex.]

---

* STOPFORD. 1930. Sensation and the Sensory Pathway. (London, Longmans.)

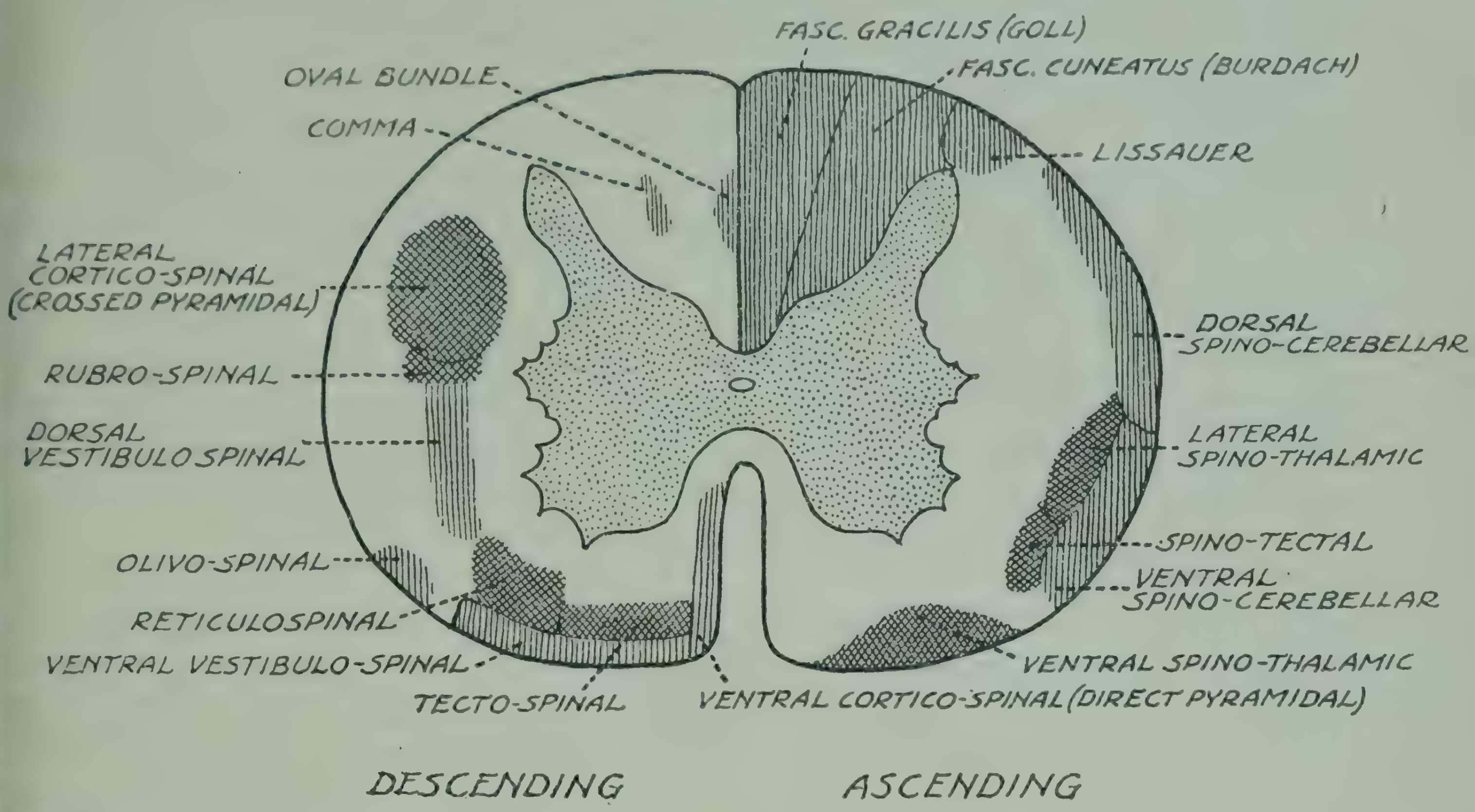

FASC. GRACILIS (GOLL)
OVAL BUNDLE
FASC. CUNEATUS (BURDACH)
COMMA
LISSAUER
LATERAL CORTICO-SPINAL (CROSSED PYRAMIDAL)
DORSAL SPINO-CEREBELLAR
RUBRO-SPINAL
DORSAL VESTIBULOSPINAL
LATERAL SPINO-THALAMIC
SPINO-TECTAL
OLIVO-SPINAL
VENTRAL SPINO-CEREBELLAR
RETICULOSPINAL
VENTRAL VESTIBULO-SPINAL
VENTRAL SPINO-THALAMIC
TECTO-SPINAL
VENTRAL CORTICO-SPINAL (DIRECT PYRAMIDAL)
DESCENDING
ASCENDING

FIG. 136. Diagram of the chief Long Tracts of the Spinal Cord. Descending tracts on left, ascending ones on right. (Single hatching, homolateral, cross-hatching heterolateral. The ventral spino-cerebellar is now regarded as mainly heterolateral.)

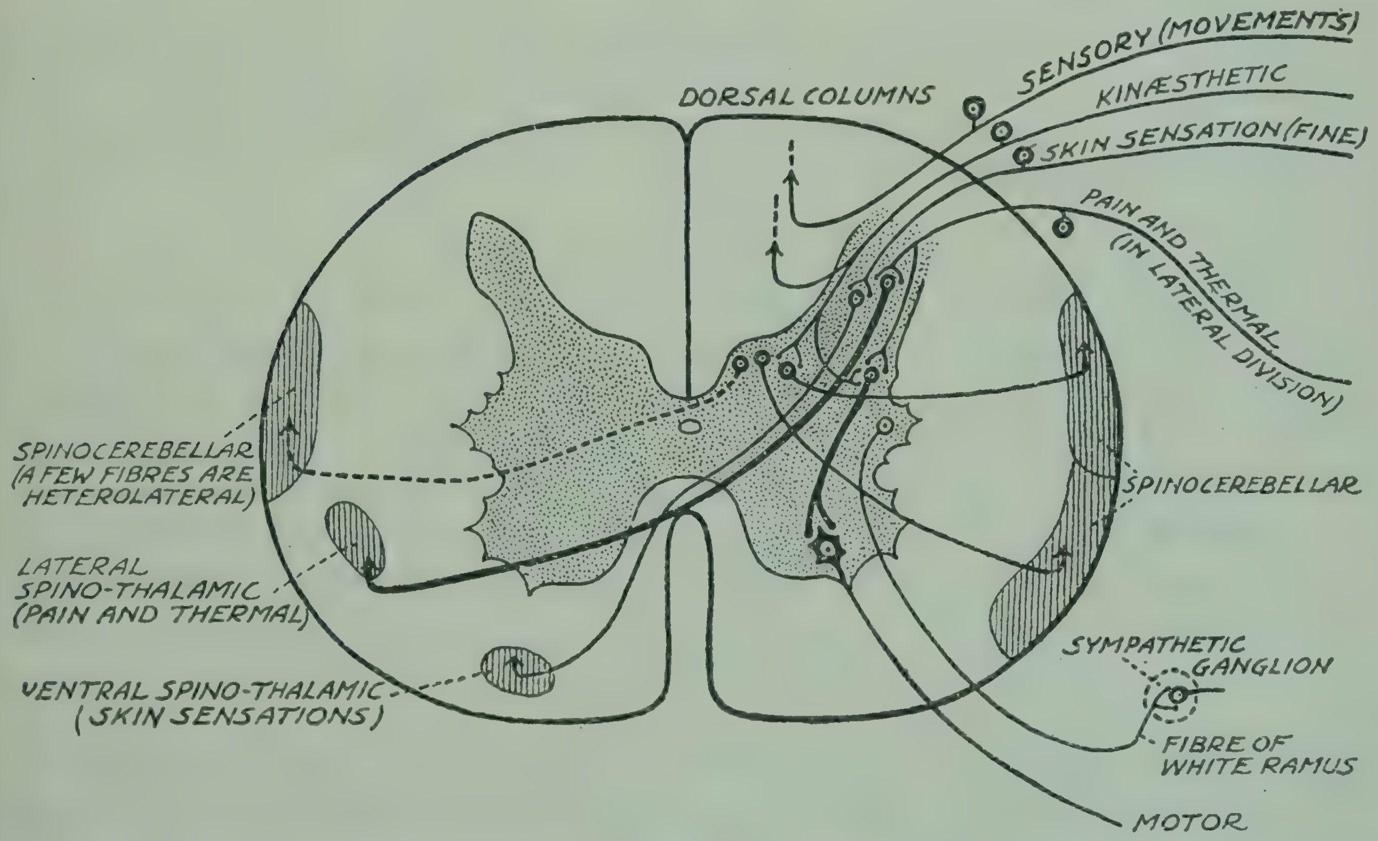

SENSORY (MOVEMENTS)
DORSAL COLUMNS
KINÆSTHETIC
SKIN SENSATION (FINE)
PAIN AND THERMAL (IN LATERAL DIVISION)
SPINOCEREBELLAR (A FEW FIBRES ARE HETEROLATERAL)
SPINOCEREBELLAR
LATERAL SPINO-THALAMIC (PAIN AND THERMAL)
SYMPATHETIC GANGLION
VENTRAL SPINO-THALAMIC (SKIN SENSATIONS)
FIBRE OF WHITE RAMUS
MOTOR

FIG. 137. Diagram illustrating the origin of the chief Afferent Tracts of the Spinal Cord.

(*b*) The *unconscious kinæsthetic impulses*; (i.) those from the skin enter the posterior root, and ascend in the columns of *Goll and Burdach* to the *nuclei gracilis* and *cuneatus*. [Second order fibres originate from the cells in these nuclei to travel as superficial or deep arcuate fibres to the inferior cerebellar peduncle by which they reach the vermis. Third order fibres may then convey the impulses to the cerebellar cortex.] The impulses along the dorsal columns travel rather slowly (70 m. p. sec.).

(ii.) The *spino-cerebellar tracts* also convey kinæsthetic impulses that do not enter consciousness. Impulses are set up by stimulation of muscle afferents (as by stretch) but not by stimulation of skin. They travel at high speed along large fibres, and most of them reach the cerebellum of the same side, where the arrival of impulses can be detected. They enter the posterior root and go to Clarke's column; from the cells in this column second order fibres ascend in both the spino-cerebellar tracts, chiefly of the same side. [Those of the *dorsal* tract entering by the inferior cerebellar peduncle, go to the vermis, from which third order fibres convey the impulses to the cerebellar cortex; some of the axons from the Clarke's column cells cross to the spino-cerebellar tracts of the opposite side. The *ventral* spino-cerebellar tract may arise from cells in Clarke's column, or (the crossed fibres) from cells scattered along the margin of the ventral horn.* At the level of the mid-brain the tract turns back abruptly to enter the superior cerebellar peduncle, where the fibres reach the vermis on the same side as that on which they started. From this, third order fibres travel to the cerebellar cortex.]

(3) PAIN AND THERMAL SENSE. The sensory fibres for pain and temperature enter by the *lateral* divisions of the dorsal nerve roots. Many of the fibres are non-myelinated.† They ascend a short distance in Lissauer's tract and relay in the cells of the *substantia gelatinosa Rolandi*. Second order fibres originating in these cells cross near the central canal to form the *lateral spino-thalamic* tract where they are dermatomally arranged, cervical nearer mid-line, sacral most lateral. [In this they ascend by the spinal lemniscus to the optic thalamus, where those concerned with the cruder sensations end. Some others for less crude forms relay round third order cells in the thalamus, which convey the impulses to the cortex.]

(4) SHORT PATH DESCENDING fibres derived from the descending axon branches include: a triangular bundle found in the lower lumbar region; the oval bundle of Flechsig, found in the lumbar region; and the comma tract, found in the thoracic region.

(5) AUTONOMIC afferent paths.

II. **Efferent Paths.** The chief projections from the cortex can be divided into the pyramidal or cortico-spinal tracts, and the extrapyramidal. The fibres of the former run *via* the medullary pyramids and end in the cord: those of the latter have a different course and destination.

(1) *The cortico-spinal tracts* arise from various sources, perhaps not all of them cortical, but at least 50 per cent. are, and this includes the large fibres which make up about 3 per cent. of the whole, and which originate in the Betz cells of area 4γ of the precentral cortex. This area also contributes many small fibres, as also does area 6 and parts of the parietal lobe (areas 3, 1, 2, 5 and in monkey 7).‡ Some fibres may be of subcortical origin. The tract is motor, and concerned with co-ordinated movements. The fibres, many of which are unmedullated, pass down through internal capsule, crura and pons to reach the medulla. Here the majority of the fibres cross, forming the *motor decussation*; part descend uncrossed. The crossed fibres form the *lateral cortico-spinal* or crossed pyramidal tract; they do not usually end at ventral horn cells directly but indirectly, by first synapsing with a cell in the dorsal horn and then perhaps again with one in the intermediate column; from here the impulse passes to the ventral horn cells; the axons of the ventral horn cells convey the impulses to the muscles. Those which do not cross travel in the *anterior cortico-spinal* or direct pyramidal tract, but most of them cross near their terminations to

---

* COOPER and SHERRINGTON. *Brain*, 1940, **63**, 123.

† RANSON. *Arch. Neurol. Psych.*, 1931, **26**, 1122.

‡ WOOLSEY and CHANG. *Fed. Proc.*, 1947, **6**, 230.

reach dorsal horn cells of the opposite side, from which fibres convey the impulses to ventral horn cells and so to the muscles. Thus, nearly all fibres from the right motor cortex are distributed to the left side of the body. The cortical cell and its axon leading to the spinal neurones is called the "*upper motor neurone*." It is of slow conduction rate (18–60 m. in cat), and its primary action is on flexor-reflex arcs, operated through interneurones in the dorsal horn.

In the lower mammals, impulses for posture and group movements probably travel largely in the *rubro-spinal* tract. This arises in the large-celled portion of the red nucleus. The fibres cross to the opposite side, and descend till they finish at anterior horn cells. These give off motor fibres to the muscles.

By means of the rubro-spinal tract the final common motor path is put into communication with the cerebellum of the *same* side (because this cerebellar path crosses twice, once in the superior cerebellar peduncles, and once in the rubro-spinal tract decussation). It also establishes connection between the corpus striatum and the final common paths of the opposite side. The cerebellum and corpus striatum also exert an influence on the final common path through the reticulo-spinal tract.

(2) BULBOSPINAL CORRELATION SYSTEM. The tracts of this system are ventrolateral in position, and represent a spinal continuation of the dorsal longitudinal bundle, with important contributions from the reticular and vestibular nuclei. The reticulo-spinal and vestibulo-spinal tracts play an important part in the control of posture.

(i.) The *reticulo-spinal tract.* This arises from the reticular formation of the brain stem and ends at the ventral horn interneurones.

The conduction is rapid, and the fibres synapse with interneurones from which the motoneurones are affected (or run direct to the latter). The reticulo-spinal tract is also directly activated on stimulation of the motor cortex, and as conduction is rapid, impulses reach the muscles earlier than those arriving *via* the pyramidal tracts, and so, perhaps, prepare the way for voluntary movement.

(ii.) The *vestibulo-spinal tract.* Impulses travel from the hair cells of the vestibular organs, and from other sources, to *Deiter's nucleus.* The cells in this nucleus give off fibres which descend by the *vestibulo-spinal* tract to the ventral horn interneurones in the cord and so to the motoneurones.

(iii.) The *tecto-spinal tract.* As the result of light falling on the retina, impulses are transmitted to the *superior corpora quadrigemina* (*colliculus superior*). Thence, fibres proceed downwards and cross at Meynert's tegmental decussation, forming the *tecto-spinal tract* by which they reach the anterior horn cells, and so reach the muscles. The connections, through the dorsal longitudinal bundle, facilitate co-ordination between the nuclei of the IIIrd, IVth, VIth and VIIth cranial nerves, and the eye and neck movements are harmonised.

(3) SYMPATHETIC EFFERENTS, *e.g.* (*a*) the pupil-dilating fibres. These fibres probably originate in or near to the superior corpora quadrigemina. They descend as far as the eighth cervical or first dorsal nerve roots. Here they probably relay at lateral horn cells, from which white rami proceed out by the ventral roots to the sympathetic chain. In this they pass up the cervical sympathetic to the superior cervical ganglion, where they synapse with cells from which postganglionic fibres proceed to the eyeballs, and so to the dilator muscle of the pupil.

(*b*) The vaso-motor fibres. These travel down the cord from the vaso-motor centre in the medulla by an unknown route. They relay in the lateral horn cells. Fibres from these cells pass out at the ventral roots of Th 1 to L 2, forming part of the white rami communicantes to the sympathetic.

There are many other sympathetic efferents (*v.* Chapter XXIII.)

EFFECTS OF SECTION OF CORD AND ROOTS

Complete Section of the Cord. Complete section of the cord results in permanent loss of voluntary movement of all muscles below the lesion. Thus, a monkey which has suffered a section of its lower cervical region will have paralysis of the lower limbs and part of the trunk, while parts above the section are apparently quite normal. The animal may contentedly direct its gaze through the window or may amuse itself by catching flies on the pane. There is also complete loss of all true sensations in parts of the body below the lesion ; reflexes such as micturition and defæcation become automatic, after a period of paralysis; and there is temporary vaso-dilatation below the lesion. These effects will be described later.

Hemi-Section of the Cord. The symptoms that follow division of the right or left side of the cord were first described by Brown-Séquard, and, after a preliminary period of spinal shock, are: (1) complete paralysis of muscles below the lesion on the same side; (2) increased postural tone and reflexes in muscles below the lesion on the same side; (3) pain, temperature, and some sensibility to touch is lost on the *opposite* side below the lesion; (4) kinæsthetic sensations are lost on the same side below the lesion; (5) temporary vaso-motor paralysis, and therefore a higher temperature below the lesion on the same side. In addition, as a result of irritation to the cut sensory fibres, there may be pain referred to parts of the body below the section. It is referred to the *opposite* side of the body to that on which the lesion of the cord has occurred. If the lesion has destroyed a length of the spinal cord, then in addition to the above we shall have the symptoms of damage to the ventral and dorsal nerve roots.

Section of the Dorsal Roots causes (most clearly when several consecutive roots are cut): (1) loss of all sensations from the limbs or other parts supplied; (2) weakness and inco-ordinated movements of the muscles owing to inability to receive kinæsthetic sensations from them; (3) injuries to the skin of the limbs supplied, owing to the absence of protective sensory impulses; (4) complete loss of tone in muscles of limbs supplied, since the kinæsthetic impulses which are responsible for them do not reach the anterior horn cells. If the section has been between the posterior root ganglion and the spinal cord, the degenerative changes affect the fibres entering the cord. These results are found in the disease tabes dorsalis. If, on the other hand, the lesion has affected the fibres between the posterior nerve roots and the sensory end organs, the fibres degenerate between the cut and the sensory end organs (see Fig. 138).

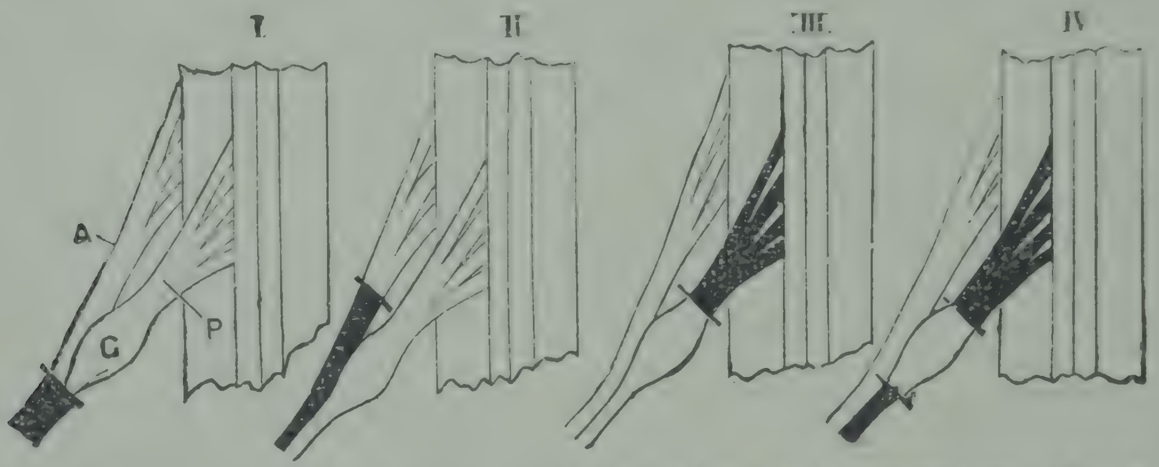


Fig. 138. To illustrate the extent and direction of degeneration (shown black) as a result of Section of the Spinal Roots. (From Yeo.)

Division of: I, Whole nerve trunk, below ganglion; II, anterior root; III, posterior root above ganglion; IV, posterior root above and below ganglion.

The above statement regarding the direction of degeneration of fibres after section of the dorsal roots is only an approximate one. Actually, many fibres remain undegenerated in the central stump, after section proximal to the ganglion,* and it has also been shown that efferent impulses, detectable electrically, pass outwards along the dorsal roots.† One among several explanations is that given by Barron and Matthews,‡ viz. that fibres enter the cord by one rootlet, and, after passing a short distance up the dorsal columns, a branch emerges by another rootlet, and passes to a second receptor at the periphery: there is the usual ganglion cell on the course of the normal afferent fibre, but none on the recurrent collateral.

Section of the Ventral Roots (several consecutive ones) causes in the part supplied: (1) complete loss of tone in muscles; (2) complete paralysis of muscles; (3) temporary vaso-dilatation owing to cutting of fibres which emerge as the white rami communicates to the sympathetic system; (4) nutritional disturbances, *e.g.* disuse atrophy of muscles, etc., unless measures be taken to prevent this: (5) changes in excitability and in chronaxie,

---

* Hinsey. *J. Comp. Neurol.*, 1934, **59**, 117.

† Barron and Matthews. *J. Phsyiol.*, 1935, **85**, 73.

‡ Barron and Matthews. *J. Phsyiol.*, 1935, **85**, 104.

(v. p. 194); (6) degenerative changes in the nerve fibres between the cut and the periphery.

FUNCTIONAL LEVELS IN THE CENTRAL NERVOUS SYSTEM

Much of the investigation of the functions of the central nervous system involves the operative cross-section of the nervous axis at particular levels, or the removal of portions of the central nervous system. Thus a section may be made at any required level of the spinal cord, medulla or mid-brain; or part or the whole of the cerebral hemispheres, the cerebral cortex, the cerebellum, etc., may be extirpated. Animals so treated represent fundamental preparations for the study in simplified form of central nervous functions.

When the spinal cord is cut in the cervical region we have the *spinal preparation* *; when the section is made at the mid-brain we have the so-called *decerebrate preparation*; if only cerebral cortex is removed, a *decorticate preparation*, and so on.

An important principle of neurology was first enunciated by Hughlings Jackson; "increasing differentiation is attended or followed by more complete integration," and the means by which this integration is effected is a problem of fundamental interest. When a lower part of the central nervous system is separated by section from a higher part, two consequences ensue. First there is *loss of certain functions* of a higher order by reason of interruption of the paths by which the lower centres are controlled from the higher ones. So, for instance, on section at the mid-brain, there is loss of voluntary movement in the trunk and limbs. Secondly, certain new properties may appear as a result of what is known as "*release of function.*" For instance, after mid-brain section there is an exaggeration of the tonus of muscles, particularly extensor muscles, normally employed in the maintenance of posture. This result appears to be due to the removal of a restraint, normally exercised by the higher centres on centres in the medulla, thus revealing or releasing a reflex exaggeration of tone in the extensor muscles (decerebrate rigidity). When the higher centres are removed, the lower ones are released from their control, as first suggested by Hughlings Jackson in 1884. It is probable that primitive behaviour responses of the lower and phylogenetically older parts of the central nervous system are normally largely dominated and repressed by the higher levels of the nervous system.

The Spinal Animal. In the spinal animal proper, section is made at the junction between the spinal cord and the medulla, and often the entire head is removed, the body and limbs then being known as a "*decapitate*" *or* "*spinal*" *preparation* (Fig. 186, p. 275). Since breathing ceases, artificial respiration by means of a pump is carried out, and the preparation then remains alive for many hours, and provides a valuable object for acute laboratory experiments.

When section is made at lower levels, we speak of an animal with *spinal transection*; if section is above the fifth cervical segment, however, artificial respiration will have to be maintained if the animal is to survive for more than a few minutes, and this is inconvenient. It is customary, therefore, where possible, for spinal transections to be made at lower levels, so that the cervical cord, which gives origin to the phrenic nerves, will remain in connection with the higher parts of the nervous system from which respiration is controlled. Such animals, or human subjects with lesions, *e.g.* by bullet wounds, etc., in the dorsal or lumbar cord, may survive for months or years, and provide objects in which the reactions of the spinal cord may be studied without the interfering influences from higher parts.

For a variable length of time after the transection of the cord, a condition of depression prevails at levels supplied by the cord below the section. This state of depression is called *spinal shock*. In lower vertebrates such as the frog, it may pass off in a few minutes, but in higher animals it may last for hours, days, or, in man, weeks. During spinal shock there is paralysis of many reflexes, the limbs are flaccid, the blood pressure low, and the bladder fails to empty. All extensor reflexes are depressed.* The flexor reflex and the knee-jerk, however, are usually retained in lower mammals, though in man the knee-jerk returns late and is then feeble.

---

* The term "preparation" in physiology means a surviving portion of an animal.

Cause of Spinal Shock. Spinal shock was formerly thought to be a result of the irritation due to the section, but is now recognised as a release phenomenon, due to the cessation of a stream of impulses from above, which normally have a facilitating action.† The reasons for this belief are (1) that if, after the shock has passed off a second section is made lower down, no shock results, (2) that if a ring of the upper spinal cord be cooled, so that conductivity is abolished at that level, shock develops although no irritation is caused by the cooling, and when the cooling agency is removed the shock vanishes again,‡ (3) that, if the cord is only partly divided, evidence of shock only appears when the vestibulo-spinal and reticulo-spinal tracts (in cats) § or the cortico-spinal tracts (in man) are sectioned.

Alteration in the reflex excitability of the spinal cord is not entirely confined to the parts of the cord below the section, since the parts above the section often show an increased excitability of reflexes (the Schiff-Sherrington phenomenon). This may be attributed to release from an inhibitory influence of the lower segments of the cord on the upper segments.

Spinal Reflexes. The *flexor reflex* is obtained when a *nociceptive stimulus* (*i.e.* a stimulus which in a conscious animal would cause pain) is applied to a hind foot or leg, and consists of a flexion of the same (or ipsilateral) limb at all joints (Fig. 139). The *knee-jerk* is obtained by tapping the patellar tendon while the quadriceps extensors are put slightly on the stretch by the

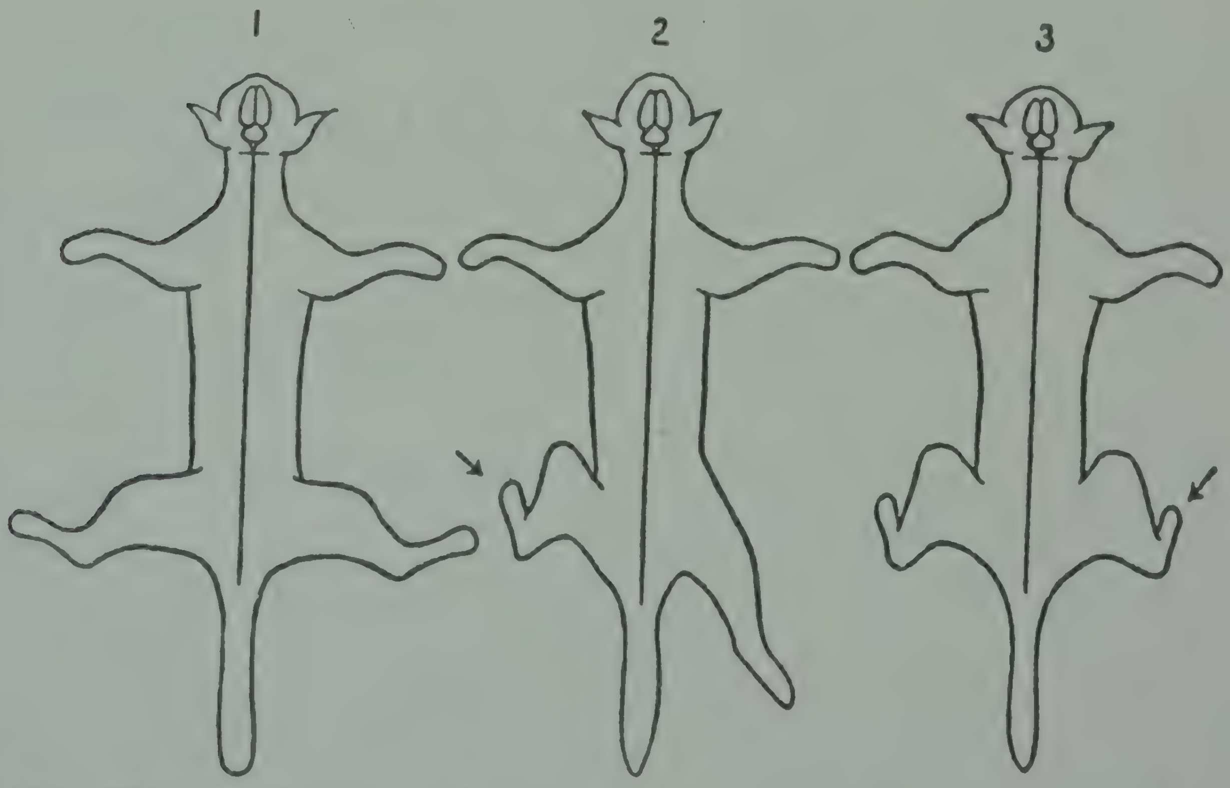


FIG. 139. Diagram illustrating Reflexes in the Spinal Cat. (1) The initial pose of the spinal animal. (2) Stimulation of left hind foot causes ipsilateral flexion and contralateral extension in the hind limbs. (3) On simultaneous stimulation of both feet there is flexion of both legs, since flexion reflex is prepotent. (SHERRINGTON. "Integrative Action of the Nervous System." Constable, 1906.)

---

* In acute experiments the extensor reflexes emerge earlier if an intravenous injection of ephedrine (1 mg./kg. for dogs) is given. HINSEY, RANSON and ZEISS. *J. Comp. Neurol.*, 1931, **53**, 401.

† LIDDELL. *Brain*, 1934, **57**, 386.

‡ TRENDELENBURG. *Pflüger's Arch.*, 1910, **135**, 469.

§ FULTON, LIDDELL and RIOCH. *Brain*, 1930, **53**, 311, 327.

weight of the leg, and consists of a strong extension of the leg at the knee joint.

As the condition of shock wears off, other reflexes make their appearance. Among these are the extensor reflexes. The *crossed extensor reflex*, *i.e.* extension in the opposite (contralateral) hind limb is obtained, together with the ipsilateral flexion reflex, when a nociceptive stimulus is applied to the foot or to any afferent nerve from the leg (Fig. 139 (2)). This is clearly a co-ordinated response, enabling the animal to throw more weight on one leg when the other is injured and therefore flexed out of the way. Various *stretch reflexes* also return, but these will be discussed later.

The *extensor thrust reflex* in spinal animals is only seen after recovery has made considerable progress. It consists in a sharp extensor push of the leg when firm pressure is applied to the pad of the foot. Flexion of the opposite leg accompanies it, and the two actions represent a compound reflex which forms the basis of galloping.

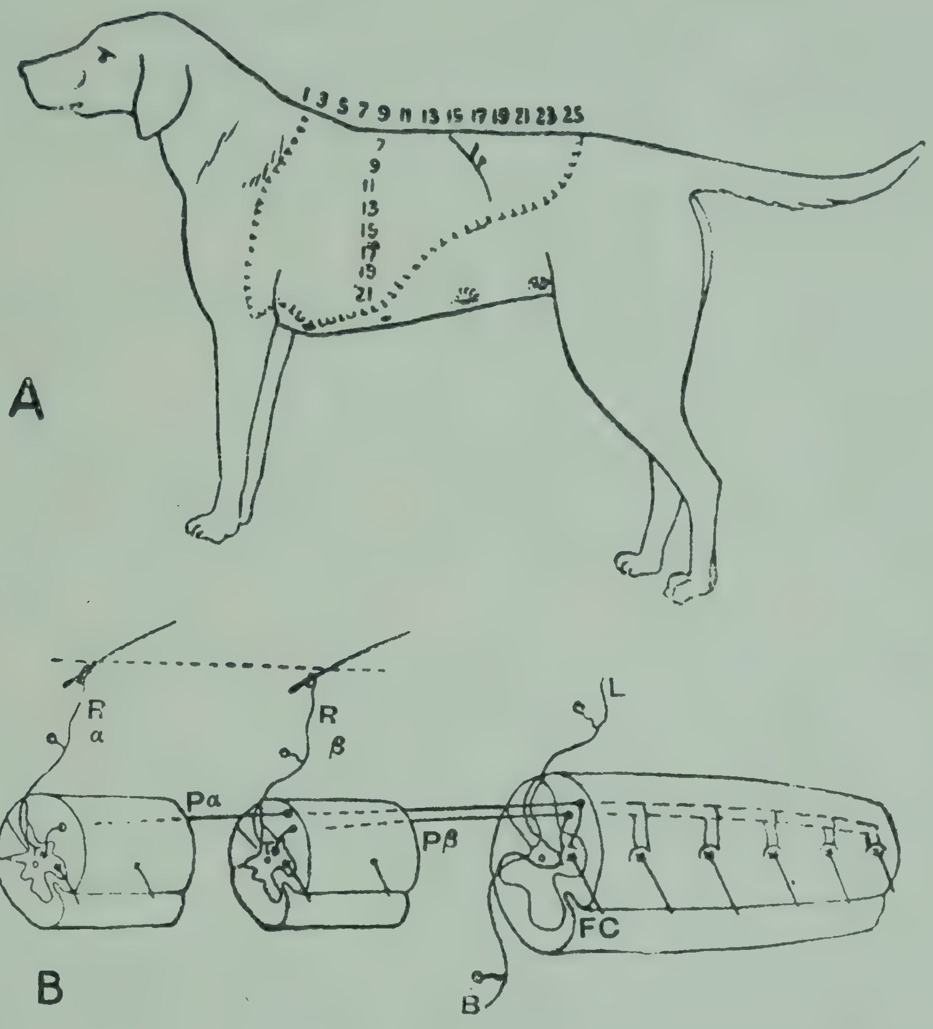


FIG. 140. A, the area from which the scratch-reflex can be evoked. B, the reflex arcs involved. (SHERRINGTON. "Integrative Action of the Nervous System." Constable, 1906.)

Another still more complex reflex late in returning is the *scratch reflex*, well seen in the spinal dog. If an appropriate stimulus is applied to a certain area (Fig. 140) of skin over the flank and shoulders, the ipsilateral hind limb executes rapid scratching movements, while the contralateral limb is rigidly extended. The appropriate stimulus should be a continuous one, slowly moved along among the hairs, in order to offset the rapid adaptation of the skin receptors, and the significance of the reflex is self-evident

Isolated Segments. For some investigations it has been found possible to isolate short portions of cord consisting of one or two segments only, by making sections above and below. In portions consisting of only two or three segments, quite normal reflex responses can be obtained.*

The Decerebrate Preparation. Division of the nervous axis at the level of the mid-brain (and destruction of the parts above the section) gives a very valuable preparation. In this, the body is left under the management of the spinal cord, medulla, cerebellum and pons, with variable parts of the mid-brain, according to the level of the section. It is sometimes called a "bulbo-spinal" preparation. As with the decapitate preparation, after the operation is completed no further anæsthetic is required, because the higher centres are all removed; unlike the decapitate preparation, various complex automatic and reflex acts are preserved. Thus respiration continues, swallowing and vomiting can take place, the arterial blood pressure remains high, and various cardiac, vascular and respiratory reflexes are retained. Temperature regulation and righting reflexes are lost, however, as in the spinal animal, but, on the other hand, there is nothing corresponding to spinal shock. It is clear, therefore, that the nervous mechanism of temperature regulation is situated higher than the level of the mid-brain, but that the cells of origin of those tracts, division of which causes spinal shock, must be somewhere between the upper spinal cord and the mid-brain. Not only does the decerebrate animal *not* show the transitory extinction of extensor reflexes characteristic of spinal shock, but instead, it shows such great exaggeration of all extensor reflexes that many of the extensor muscles are aroused to a state of permanent hypertonus. The flexor responses seen in the spinal preparation can all be elicited, however, though not so easily.

---

* TEN CATE. *Arch. néerl. Physiol.*, 1932, **17**, 331, 525.

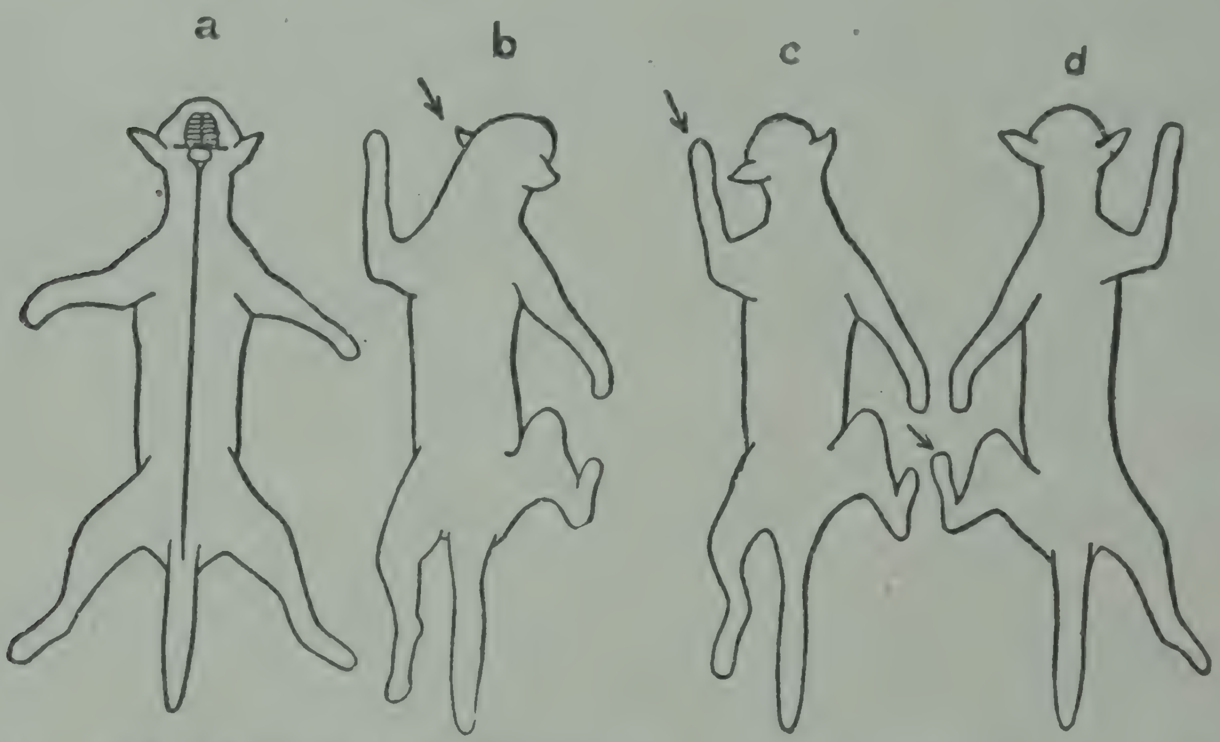


FIG. 141. Reflex Figures in Decerebrate Cat. (a) Normal decerebrate rigidity—extension of all limbs. (b) Stimulation of left pinna—flexion of left fore and right hind limbs, increased extension of others. (c), (d) Stimulation of left fore and left hind limbs respectively. (SHERRINGTON.)

The extensor tonus is evident as a rigidity of the limbs,* spoken of as *decerebrate rigidity* (Figs. 141, and 186, p. 275). It affects the anti-gravity muscles, *i.e.* those muscles involved in keeping the animal in a normal standing posture, which in most quadrupeds are the extensors at all joints of the four limbs, and the muscles that keep the head and tail from falling forwards.

If we accept spinal shock as a release phenomenon, due to the removal of an influence from above, which augments the extensor reflex responses, then we must recognise decerebrate rigidity as due to the retention of that influence. It is thus a release phenomenon of another type. We may suppose that the influence of centres in the medulla, or other higher parts that are present in the preparation, in exaggerating extensor responses is now left in undisputed possession of the field, whereas in the intact central nervous system the effects of these centres are subordinated to still higher centres. Hence a grossly exaggerated picture of reflex standing results, which under the influence of higher centres would be modified to give the normal stance.

The decerebrate preparation is accordingly very useful for the investiga-

---

* SHERRINGTON. *J. Physiol.*, 1897, **22**, 319.

tion of phenomena of extensor reflexes and their inhibition. It also shows many instances of complex reflexes involving co-operation between different segments of the cord. For instance, there is the phenomenon of *reflex figures*, *i.e.* if a pinna or one limb is stimulated characteristic flexion and extensions of limbs as shown in Fig. 141 (b).

It will be clear that these various reflex figures play their parts in the normal movements of the animal.

The Decorticate Preparation. Removal of the entire cerebral cortex gives a preparation that in many respects resembles a normal animal.* In dogs and cats, most bodily functions and also temperature regulation, locomotion and posture are normal, and *righting reflexes* are present, *i.e.* if the animal is put into an abnormal position, it rights itself. There is no extensor rigidity so long as a normal position is maintained, though if held up by the neck an extensor hypertonus resembling decerebrate rigidity makes its appearance. If sections of the brain are made coronally from before backwards, no sign of decerebrate rigidity appears until section reaches a plane passing through or just behind the posterior part of the thalamus,† thus indicating that decerebrate rigidity is due to release of lower postural centres (Deiter's nucleus), from the control of centres lying anterior to the pons. Posture and righting reflexes are, in fact, retained almost normally if section passes anywhere anterior to the mid-brain, so that righting reflexes are probably controlled from the mid-brain.

If only the cerebral *cortex* is removed, leaving behind the basal ganglia, the animal may retain many reflexes connected with special senses ; thus it will make attempts to eat articles that smell like food, moves to sounds, shows constriction of the pupil of the eye to light, etc. All of these reactions, however normal in appearance, are reflexes mediated by subcortical centres and have no connection with any actual sensation, since the animal is totally deaf, blind and unconscious.

Still more remarkable are the appearances of the expression of primitive emotions, such as rage and fear. A decorticate dog or cat, if disturbed or stimulated in almost any way, however slightly, will exhibit all the signs of violent and exaggerated rage, biting, struggling, lashing of tail, erection of hairs, high pulse-rate and blood pressure, etc. These reactions, though complex and apparently purposive, are also reflex in nature, and are not associated with any state of consciousness on the part of the animal. They are accompanied by intense outbursts of sympathetic activity. These subconscious manifestations are probably due to activity in the posterior part of the hypothalamus,‡ and are release phenomena, due to the removal of that cortical control which regulates and rationalises the behaviour of the whole animal. It is probable that unreasoning fear or rage, and emotional instability, in man, are often associated with release of the hypothalamic mechanisms from normal cortical control.

In the higher apes (and we may suppose that it would be similar in man), decortication leads to more profound and evident losses of function than is the case with lower mammals. Normal locomotion is lost and the animal lies on its side in an abnormal posture, with the undermost limbs extended and rigid and the uppermost ones flexed.§

SPINAL TRANSECTION IN MAN

The consequences of spinal transection in man were studied in great detail by Head and Riddoch ‖ on men wounded in the World War I. At the moment of injury, if the transection is complete, the subject feels that the lower part of his body has vanished. The patient will never recover, though his condition alters as time goes on. (I.) There is first a prolonged *stage of shock*, with complete loss of movement and sensation in the parts below the lesion, and a feeling of a tight or painful girdle round the body is often present just above it. The muscles are flaccid, the limbs flushed but cold, and there is no sweating below the lesion. Blood pressure is low. Nearly all the muscular reflexes vanish. The anal and bladder sphincters are tightly contracted, the detrusor flaccid, and urine and fæces tend to be retained. This stage lasts for two or three weeks.

---

* Goltz. *Pflüger's Arch.*, 1892, **51**, 570.
† Thiele. *J. Physiol.*, 1905, **32**, 358.
‡ Bard. *Amer. J. Physiol.*, 1928, **84**, 490.
§ Bieber and Fulton. *Arch. Neurol. Psychiat.*, 1938, **39**, 435.
‖ Head and Riddoch. *Brain*, 1917, **40**, 188.

(II.) In the second stage there is a *return of reflex activity*. The bladder sphincters become less tonic, and later the bladder is reflexly emptied. The arterial pressure returns to normal and there is some improvement in circulation in the legs. Tone begins to return a little in the flexor muscles, and flexor reflexes appear, the first hint of this usually being the occurrence of slight upward movement (" dorsiflexion ") of the big toe (the Babinski reflex),* when the sole of the foot is scratched. Later, the complete flexor reflex (dorsiflexion of toes and of foot, flexion of knee and hip with abduction of thigh), slowly emerges, and tone in flexor muscles is back to normal.

The Babinski reflex is a primitive one, seen in infants, which is absent from normal adults : in these the response of the toes is plantar-flexion. This is converted to the dorsiflexor response in any circumstances in which the cortico-spinal tract is severed. It is unfortunate that the name " extensor response " has been attached to the Babinski reflex. Dorsiflexion of the big toe as seen in the Babinski response, is part of a generalised nociceptive *flexor* reflex,† since the extensor longus hallucis shortens the limb ; when the response is powerful it can be elicited from a large area of the limb, and is accompanied by a crossed extensor response.

The abdominal reflex, a contraction of the recti on scratching the surface of the abdomen, is another conspicuous flexor reflex.

Extensor tone never shows much recovery, and extensor reflexes, such as the knee-jerk are weak, but sharp and brief. A weak crossed extensor response may accompany a flexor reflex. Owing to the preponderance of flexor over extensor tonus, the lower limbs usually remain semiflexed, and hence this condition is often called *paraplegia in flexion*. Flexor response to even a slight stimulus almost anywhere in the paralysed region may ultimately become very widespread and powerful, and becomes the *mass reflex*. This is due to irradiation, and is accompanied by sweating in the paralysed regions and by evacuation of the bladder owing to the rise of intra-abdominal pressure.

During this second stage, reflex micturition and defæcation make their appearance. Gradually also, there appear, at intervals, " spontaneous " reflex flexor movements of the lower limbs, which ultimately become almost continual. Hence the muscles do not waste so much as in cases of destruction of the lower motor neurone, *e.g.* that on section of the motor nerves.

In a third, and much later, stage, reflex activity may gradually fail again and then the muscles do waste. Trophic changes such as the formation of ulcers and bed-sores further complicate the patient's existence.

Incomplete transection of various degrees, of course, also occurs in man and is frequently distinguished by the fact that extensor tone and extensor reflexes may be enhanced instead of being lost. Flexor reflexes, including the abdominal reflex, on the other hand, are depressed or absent. Hence the condition is often called " *paraplegia in extension*." The difference is attributed to the section of the cortico-spinal with retention of the vestibulo-spinal and reticulo-spinal tracts, so that the state of reflex activity in the cord more resembles that of the decerebrate than of the spinal animal.

Trophic Functions of the Cord. The reflexes which are excited by nocuous stimuli must be regarded as *prepotent*, in that their inhibitory effect on other reflexes is more marked than that produced by any other quality of stimulus. In the struggle for existence, the reaction to nocuous stimuli must predominate over those due to any other kind, since it is essential for the survival of the animal that these stimuli should be removed or avoided.

---

* FULTON and KELLER. The Sign of Babinski. 1932 (Thomas, Springfield).

† WALSHE. *Brain*, 1914, **37**, 269.

It is natural, therefore, that after complete section of the spinal cord, or of the afferent nerves from any part of the surface of the body, there should be a tendency to trophic disturbances such as the formation of ulcers, etc. A spinal patient may easily be burnt by a hot-water bottle. After section of the first division of the Vth nerve, ulceration of the cornea is often produced. These effects are due to the absence of the normal protective reactions of the part, and can be prevented by scrupulous cleanliness and protection of the anæsthetic part from all possible injuries.

There are other trophic effects caused by nerve lesions which cannot be ascribed to the mere absence of protective reflexes. Thus, inflammation of the posterior root ganglia often sets up "herpes zoster," or "shingles," in the region of cutaneous distribution of the corresponding sensory nerve. Changes in the skin ("glossy skin"), nails and hair are often seen after irritative injuries of nerves to the part. Section of a motor nerve causes rapid changes in the skeletal muscles supplied, which become smaller, and after months or years may disappear altogether, being replaced by connective tissue.

It seems that the nutrition of a tissue is determined by its activity, and this in turn is under the control of some nerve path. Section of the nerve path, by cutting away the impulses which normally maintain the activity of the part, must at the same time seriously affect its nutrition. Thus, muscles such as the sphincter ani, which, though striated, are only under the reflex control of the central nervous system, only slowly undergo degeneration after section of their nerves, or after extirpation of the lower part of the spinal cord.

The Brown-Séquard Syndrome. One of the most impressive of neurological states is that in which, as a result of injury, or of pressure from a tumour, there is unilateral interruption of the spinal pathways. There follows a loss of movement on the same side as the injury, together with a qualitatively differentiated loss, or impairment, of sensation on one or other side. Proprioceptive sensations are lost on the same side as the lesion, because their fibres, in the posterior columns, do not cross until the medulla is reached. Pain and thermal sensations are lost on the opposite side, because their fibres cross (into the lateral spino-thalamic tract) soon after entry into the cord. Fine skin sensation is not lost because some fibres remain uncrossed, though most cross over.

Section of the spino-thalamic tracts has been effected surgically for relief of intolerable pain.

CHAPTER XVI

THE EXCITABILITY OF NEURONES AND THE PROPERTIES OF CONDUCTION AT SYNAPSES

Conduction in a reflex arc differs from that in peripheral nerves, in that part of it takes place through grey matter, and hence we must look to the properties of the bodies of the nerve cell (soma) or to properties of the synapse for an explanation of the differences between nerve fibre conduction and reflex arc conduction.

It is implicit in the neurone hypothesis that the impulse which emerges by the efferent fibre of a reflex arc is not the same impulse as that which enters the spinal cord by the afferent fibre : it is a new impulse generated by the cell of which the efferent fibre is the axon. Apart from the important evidences of histology, we believe that this is so because (*a*) the time that elapses between the entry of the afferent impulse and the emergence of the efferent response is longer than would be needed for direct transmission, and (*b*) the emergent impulses in response to a given train of afferent impulses, though they may, and often do, tally impulse for impulse, with the number of ingoing impulses, do not necessarily do so ; either more or fewer may emerge.

Present-day conceptions of the mechanism of reflex action are indebted chiefly to the investigations of Sherrington and his pupils, and most of their explanations, derived from the study of spinal reflexes, have received general acceptance. It is also generally agreed that an isolated and simple reflex, as diagrammatically represented in Fig. 126, is nothing more than a convenient fiction. When we come to study actual reflexes, we find them to be less simple, and hence interpretation is a matter for detailed analysis. The essential correctness of the conclusions reached by the Sherrington school is shown by the fact that their conceptions have received important confirmation and extension from recent work, which, being more definitely directed to the study of the properties of synapses and neurones, will be outlined first.

Nature of Synaptic Conduction. When we speak of the "passage" of an impulse through a synapse, we mean that the incoming impulse is extinguished and a new impulse started in the cyton of the succeeding neurone. There are two views as to the way in which this transmission is brought about, and these may be called the *electrical* theory and the *chemical*, or humoral, theory respectively.

According to the electrical theory, the action potential in the synaptic terminals of the afferent fibres excites the body or dendrites of the second neurone, acting, in fact, as an electrical stimulus. According to the humoral transmission theory, the action potential at the synaptic endings causes, in or near those endings, a liberation of a highly active substance, the transmitter substance, which chemically excites the next neurone. We have already considered such a case for neuromuscular transmission.

For the present we shall consider the matter from the electrical standpoint, because, as a first approximation, it is simpler to do so, and later, in Chapter XXIII, shall return to the consideration of the chemical theory. Probably each of these theories represents a partial view of what occurs, and the electrical and chemical changes are only different aspects of a complex event of many concomitants.

The main requirements in electrical synaptic transmission are * : (i) that the action potentials at the synaptic knobs of the pre-synaptic endings set up local electrotonic potentials at the surface of the post-synaptic cell. These resemble the local non-propagated responses in nerve fibres (p. 180) and beyond a nerve block (p. 187); (ii) these potentials can summate, and, when the threshold is reached, start off an impulse in the post-synaptic cell and so on to its axon.

---

* Eccles. *Ann. N.Y. Acad. Sci.*, 1946, **47**, 429.

The rapid build-up of synaptic potentials is helped by the fact that many synaptic endings converge on to each nerve cell. As the action potential in the pre-synaptic fibre approaches the synaptic knob, the post-synaptic membrane would become first an anodic and then a cathodic region, as shown in Fig. 142 (*a*) and (*b*). The former would locally depress excitability, and the latter would set up a locally increased excitability due to the catelectrotonic state; when this has reached a sufficient intensity and occupied a large enough area, an impulse would be started which would then pass to the post-synaptic axon.

That there is nothing inherently improbable in this hypothesis has been shown by experiments on artificial synapses or *ephapses*.* These consist of two giant non-medullated fibres touching one another, the orientation of one fibre to the other being varied in different ways. The results of such experiments have been described on p. 184. Under the conditions of the experiment the resting fibre is not sufficiently excited to start a propagated impulse, but this result can be attained if the excitability of the fibre is raised by lack of calcium, and it could well be that in the actual synapse, the number of end-knobs, and their nearness to the post-synaptic membrane, would readily allow of the catelectrotonic state of the latter reaching the requisite threshold.

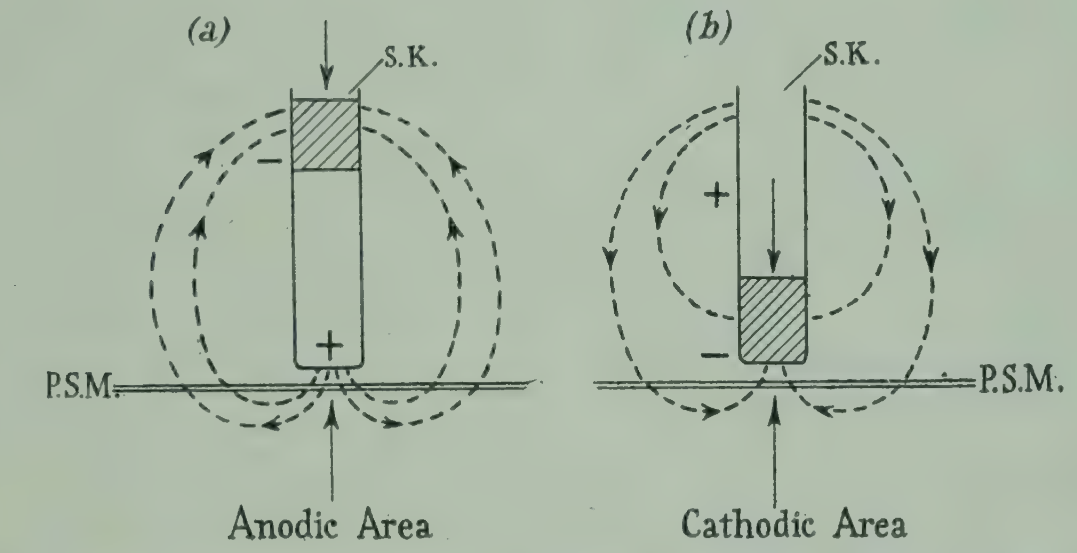


FIG. 142. Development of Synaptic Potential. S.K., Synaptic knob; P.S.M., Post-synaptic membrane. (*a*) Action potential approaching; area of anelectrotonic potential beneath knob. (*b*) Action potential arrived: area of catelectrotonic potential beneath knob. (After Eccles.)

Transmission in Sympathetic Ganglia. Synaptic transmission is reduced to its simplest terms in the sympathetic ganglia, *e.g.* the superior cervical ganglion, when stimulated by a synchronous pre-synaptic volley. The fibres which enter from below (pre-ganglionic fibres) have synapses in the ganglion, from the cells of which axons pass up (post-ganglionic fibres) in the neck and are distributed to various structures in the head. One of these structures in the cat is the plain muscle of the nictitating membrane (Fig. 143) and it is the neurones supplying this which have been especially studied. There are at least three other types of neurones in the ganglion, but for the sake of simplicity we shall only consider this one.

When a volley of impulses is started in the pre-ganglionic fibres, by stimulation at *A* (Fig. 143), a volley emerges along the post-ganglionic fibres and can be recorded from electrodes one of which is placed at *E*.† Also the response

---

* KATZ and SCHMITT. *J. Physiol.*, 1940, **97**, 471; ARVANITAKI. *J. Neurophysiol.*, 1942, **5**, 89.

† BISHOP and HEINBECKER. *Amer. J. Physiol.*, 1932, **100**, 518.

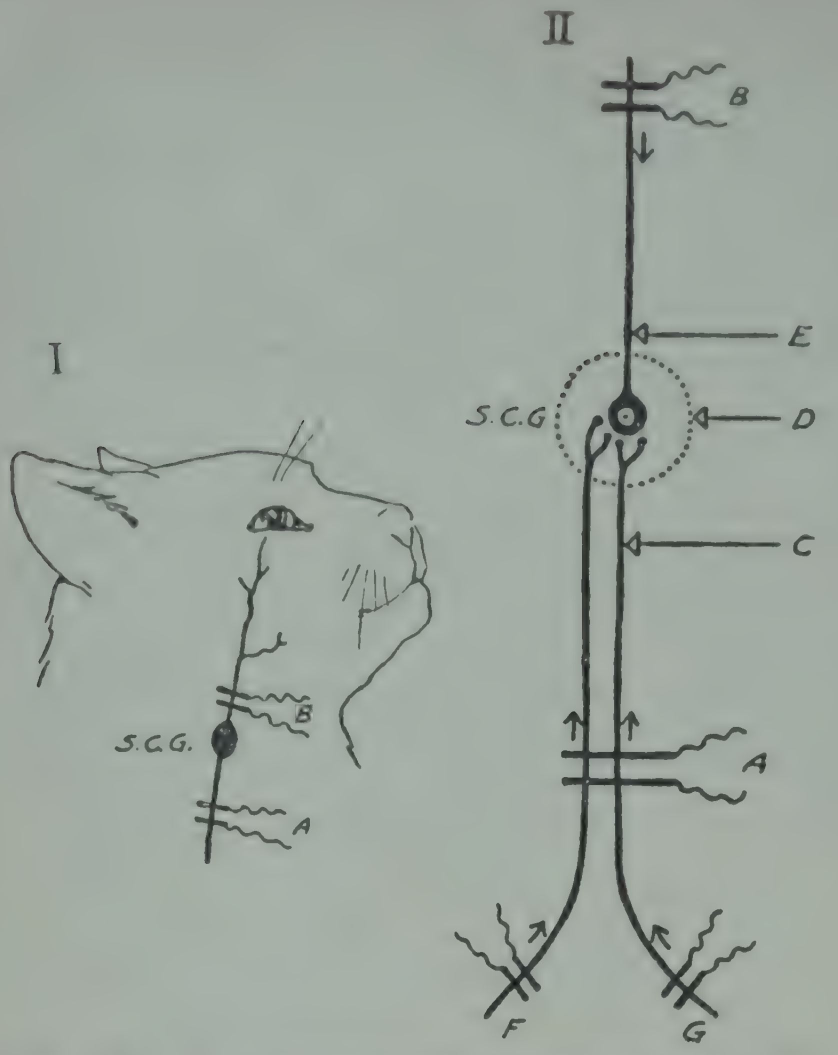


FIG. 143. I, Arrangement for stimulating the cervical sympathetic nerve, pre-ganglionically at A, or post-ganglionically at B. II, Scheme showing neurones concerned. *A*, *B*, *F*, *G*, points for stimulation. *C*, *D*, *E*, points for leading off potential changes. *S.C.G.*, Superior cervical ganglion.

can be registered by recording the contraction of the nictitating membrane.* The neurones respond to each afferent impulse with one efferent nerve impulse, and one only.† In conduction across these synapses about 3 msec. is lost (= synaptic delay). If the post-ganglionic nerve be stimulated, *e.g.* at B (anti-dromic stimulation), no impulse can be detected in the pre-ganglionic fibre, *e.g.* at electrode *C*, *i.e.* the synapse will not conduct backwards; but a disturbance is detected at the ganglion itself, showing that the antidromic impulse extends back as far as the nerve cell, and probably into any dendrites it may have. The backfiring of an anti-dromic impulse does not leave the nerve cells unaffected, but is followed by a refractory period during which a normally arriving pre-ganglionic volley produces no effect.

When the pre-ganglionic fibres are stimulated, and the action potentials thereby produced are studied in the ganglion and in the adjacent post-ganglionic fibres, these are found to be unlike those which pass along the pre-ganglionic fibre. The changes in the ganglion show three waves, viz. a spike (*S*), a negative after-

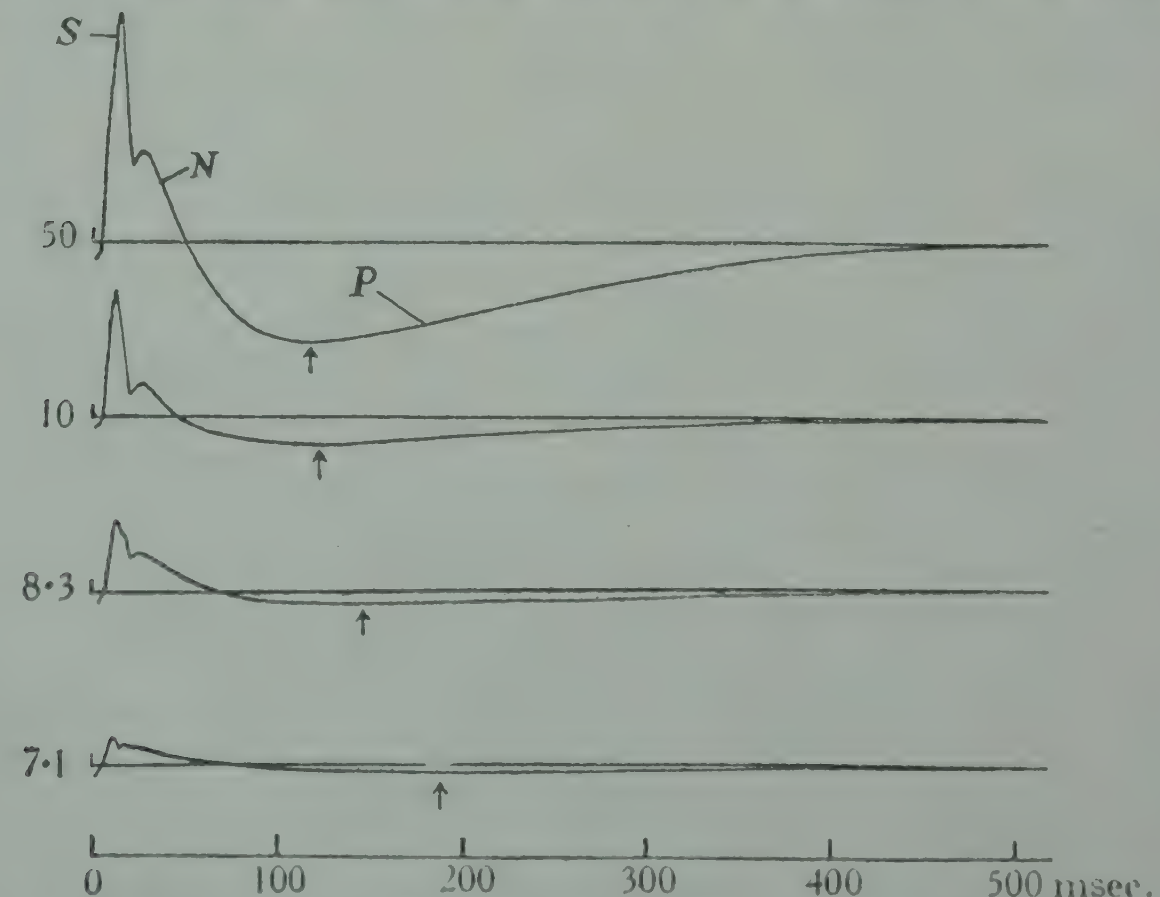


FIG. 144. Ganglionic Action Potentials in Superior Cervical Ganglion, set up by Pre-ganglionic Stimuli whose strengths are shown at left. Arrows mark the points of maximum positivity. *S*, *N*, and *P* are, respectively, spike, negative and positive waves. (ECCLES. *J. Physiol.*, 1935, **85**, 464.)

---

* BROWN. *J. Physiol.*, 1934, **81**, 228.

† BRONK, TOWER, SOLANDT and LARRABEE. *Amer. J. Physiol.*, 1938, **122**, 1.

potential (N), and a positive after-potential (P) (Fig. 144), and it was shown by Eccles * that these are related to the excitatory process in the nerve cells, in much the same manner as the spike and after-potentials of nerve fibres are related to its excitatory condition.

Immediately following the arrival of the pre-ganglionic nerve impulses at their synapses is a *synaptic delay* of about 3 msec. The spike potential of the soma then occurs if the nerve cell is to discharge an impulse along its axon. During the spike, which lasts about 5 msec., the cell is in an absolute or relative *refractory state*, *i.e.* a second impulse does not excite, or excites less readily.

The after-potentials, according to Eccles, consist of two simultaneous and opposing changes, one tending to make the ganglion (cells) negative and one positive.

At first the negative wave is the greater, but the positive one outlasts it (Fig. 145). In the algebraic resultant therefore we have first a negative potential of varying size (N) followed by a positive one (P) (Fig. 144). Both potentials increase in size as the stimulus is increased. During the prevalence of the N-wave the ganglion cells are in a state of supernormal excitability, in consequence of which they respond to a smaller second stimulus and with shorter synaptic delay than before (temporal summation). This stage may last for 150 msec. An N-wave always accompanies the application of a stimulus to the nerve cells, *e.g.* at the synapse, even if this is too weak to excite them (Fig. 144). The appearance of the N-wave coincides with the condition of the cell which was called by Eccles and Sherrington the *central excitatory state* (c.e.s.), *i.e.* a proneness to excitation. Hence if two subthreshold stimuli along the same afferent fibre are so spaced in time that the second arrives during the c.e.s. produced by the first, the raised c.e.s. may enable the cell to respond and discharge an impulse. Since the supernormal phase follows the refractory period of the cell and fibre, therefore, *temporal summation* can occur in ganglia, even when the intervals run into many msecs. When the c.e.s. reaches a certain level the cell discharges its impulse at the slightest provocation, and when it has done so this hastens the subsidence of the c.e.s.

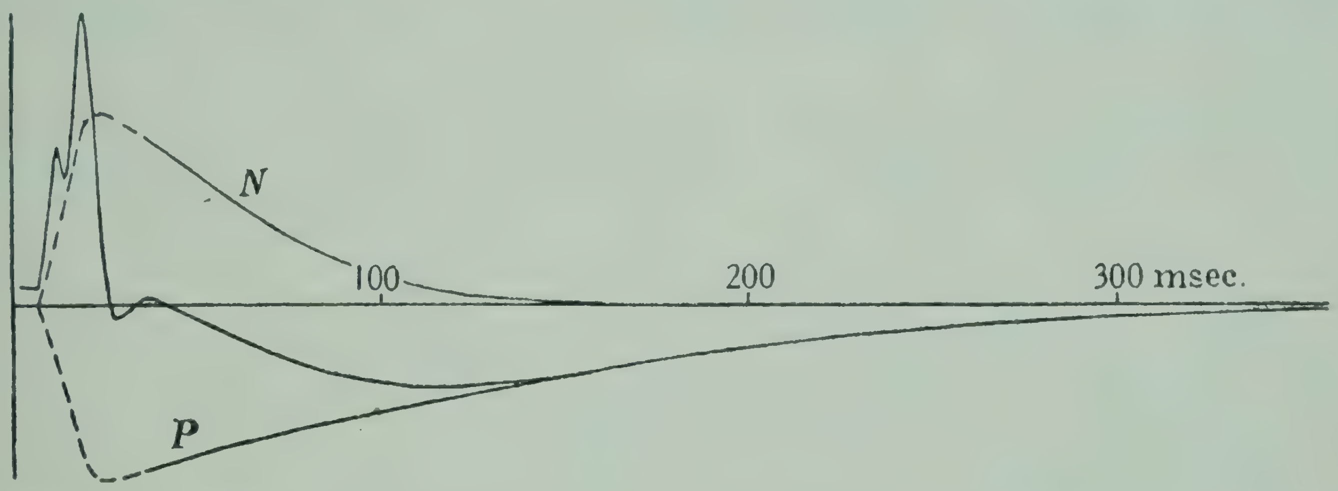


FIG. 145. Ganglion Potential Curve, and the Reconstructed Negative (N) and Positive Waves, which sum algebraically with the spike to give the complete curve. (ECCLES. *J. Physiol.*, 1935, **85.**)

The presence of *small* amounts of adrenaline may facilitate the synaptic response.†

Another kind of summation may be illustrated with the superior cervical ganglion. If stimuli of subthreshold strength be applied to two branches (of the annulus of Vieussens) at F and G (Fig. 143 II), provided the stimuli are not more than about 2 msec. apart, excitation of the cells may result. This is called *spatial summation* because the stimuli act at different points on the soma.

The P-wave, which follows the N, is indicative of a corresponding change in the excitability of the cell, which is now reduced—a condition called *central inhibitory state* (c.i.s.). During this phase, which lasts for 500 msec., the cell is less excitable, *i.e.* its threshold of excitability is raised.

The results of backfiring an antidromic impulse from the post-ganglionic fibres into the ganglion show that the neurone is caused to go through processes very similar to those which occur when it is stimulated in a forward direction.‡ Thus the ganglion shows a spike potential followed by the N- and P-waves, the P-wave being very large and prolonged. After the backfiring there is an absolute refractory period of 2 msec. followed by a long lasting state of depressed excitability running parallel with a prolonged P-wave. In other words, the cell has discharged an impulse, though this has been unable to pass backwards through the synapse, and the discharge has been followed by a prolonged central inhibitory state.

---

* ECCLES. *J. Physiol.*, 1935, **85,** 179, 207, 464; 1937, **91,** 1.

† BÜLBRING. *J. Physiol.*, 1944, **103,** 55.

‡ ECCLES. *J. Physiol.*, 1936, **88,** 1.

The after-potentials are essentially due to changes of polarisation in the cell body and do not extend far down the post-ganglionic axons; probably they merely spread electrotonically. They indicate a *state of excitability* of the cell and its synapse and *not* the propagation of a nervous impulse. The N-wave is reduced by nicotine, which also destroys c.e.s.

In further investigations Eccles,* using the stellate ganglion, showed that when this is curarised, so as to block synaptic transmission, preganglionic stimulation causes a local negative potential of the ganglion cells relative to their axons; this is due to a *synaptic potential*, which spreads decrementally and slowly decays. It is this catelectrotonic potential (resembling end-plate potential in curarised muscle), in its earlier and intense phase (peak at 3 msec.), which, in absence of curare, is able to excite the soma. The synaptic potentials show summation with repetition of stimulation. The speed of the synaptic response is one of the difficulties in the way of regarding transmission as a chemical mechanism such as by acetyl choline.† Moreover, the duration of this peak response is not increased by eserine. The c.e.s. set up in the cell does not begin until after the synaptic delay and lasts much longer, reaching a maximum at 10 to 15 msec. It can therefore have no direct action in discharging the impulse, but merely increases the "explosiveness" of the cells.

Transmission at Motoneurones. We may now consider the phenomena of excitation and transmission in less simple neurone arrangements, such as those found in the central nervous system proper. Most valuable investigations have been made by Lorente de Nó ‡ on the motoneurones of the oculomotor nucleus, which supplies the motor fibres of the third, fourth and sixth cranial nerves.

The oculomotor neurones are a final common path in three reflex arcs, one from the labyrinth, one from the fifth nerve and one from the neck muscles.

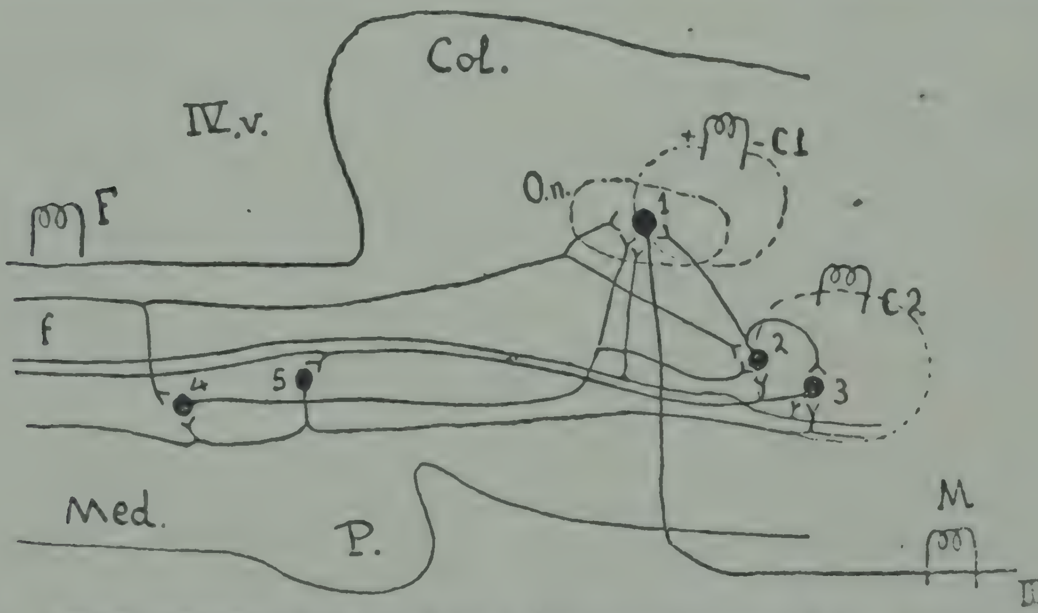


FIG. 146. Arrangements for studying the responses of the oculomotor motoneurones. IV.v. Fourth ventricle; Col, colliculi; Med, medulla; P, pons; F, C_1, C_2 and M., positions for stimulating electrodes; O.n. oculomotor nucleus; III, oculomotor nerve; f, fibres in posterior longitudinal bundle stimulated by F shock; 1, motoneurones of third nucleus; 2, 3, 4 and 5, internuncial neurones of reticular substance; Shocks at C_1 and C_2 stimulate motoneurones and internuncial cells respectively. (LORENTE DE NÓ. *J. Cell. Comp. Physiol.*, 1935, **7**, 47.)

The afferent fibres from these pass in or near to the posterior longitudinal bundle which runs near the floor of the fourth ventricle and can be stimulated there.

The arrangement of some of these experiments is illustrated by Fig. 146.

These neurones can be stimulated, either *viâ* synapses which are formed with fibres of the posterior longitudinal bundle, by electrodes in the floor of the fourth ventricle at the point F in Fig. 146, or directly at C_1, by electrodes buried in the colliculi, and the response can be recorded as an action potential in the third or fourth nerves, or in the eye muscles they supply.

---

* ECCLES. *J. Physiol.*, 1943, **101**, 465; 1944, **103**, 27.

† BROWN. *Physiol. Rev.*, 1937, **17**, 485.

‡ LORENTE DE NÓ. *Amer. J. Physiol.*, 1935, **111**, 272, 283; **113**, 505, 524. *J. Cell. Comp. Physiol.*, 1935, **7**, 47; *J. Neurophysiol.*, 1938, **1**, 187, 195, 207; 1939, **2**, 402.

By stimulation at the electrodes M, on the third nerve, the cells (1) of the nucleus receive an antidromic impulse, and, of course, the muscle is stimulated by the forward impulse. The cells and fibres concerned in these experiments are of higher excitability and faster action than those we have just studied in the sympathetic system, but apart from this difference in the time-scale there is a general similarity between the phenomena in the two places.

Direct stimulation of the nerve cells (position C_1) with one shock, gives, with no appreciable latency, a single volley of impulses; two subliminal shocks, if in sufficiently quick succession, also excite by latent addition. For this to occur they must be not more than 0·2 to 0·3 msec. apart. This shows that the excitatory process caused by stimulation of the motoneurone is of the same duration as that of nerve fibres. Having once responded, the motoneurones pass through a *refractory phase,* which is absolute for about 0·5 msec., which again is identical with that for nerve fibres.

The refractory phase is followed by a long period (30 msec. or more) of *subnormal excitability* which resembles that of a nerve fibre that has received a number of stimuli in quick succession (Fig. 97, p. 179). It is doubtful whether there is a supernormal phase resembling that shown by nerve after a single conditioning stimulus, or by the cells of sympathetic ganglia.

Thus far, then, there is a close similarity between the properties of the soma and those of the nerve fibre.

Synaptic stimulation of the motoneurones can be brought about by a single stimulus applied at the F electrodes (Figs. 146 and 147). When excited in this way there is a delay of about 0·6 msec. in transmission across the synapse. If the motoneurone does not respond within this time it will not respond at all unless restimulated by a new impulse. Since the major part of the spike potential in a nerve fibre (or synaptic knob) lasts about 0·5 msec., it would seem that it is the action potential at the synaptic knobs which isthe synaptic transmitter of excitation.

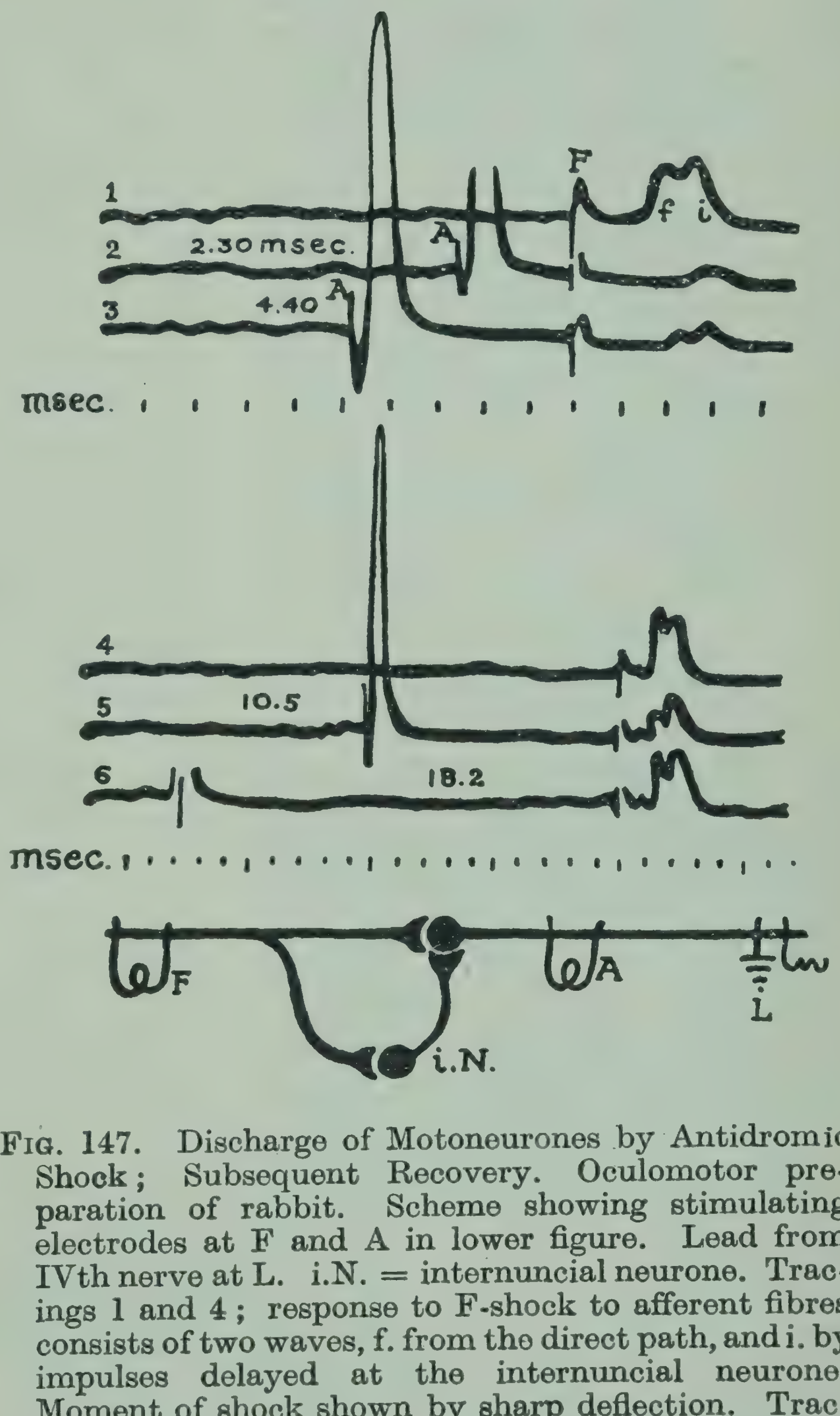


Fig. 147. Discharge of Motoneurones by Antidromic Shock; Subsequent Recovery. Oculomotor preparation of rabbit. Scheme showing stimulating electrodes at F and A in lower figure. Lead from IVth nerve at L. i.N. = internuncial neurone. Tracings 1 and 4; response to F-shock to afferent fibres consists of two waves, f. from the direct path, and i. by impulses delayed at the internuncial neurone. Moment of shock shown by sharp deflection. Tracings 2, 3, 5 and 6, the F-shock has been preceded by an antidromic maximal stimulus at A, the intervals between A and F being shown on records. Time scale in msec. is quicker for 1, 2 and 3 as shown in scales. (Modified from Lorente de Nó and Graham, *Amer. J. Physiol.*, 1938, **123**, 388.)

An antidromic impulse fired into the cells from the point A (Fig. 147) on the motor nerve, discharges the neurones and is followed first by a refractory phase and then by a state of depression (Fig. 147) which lasts for at least

30 msec., just as happens when the cells are stimulated directly; during these stages of depression synaptic stimulation (by an F-shock) produces less or no effect and synaptic delay may be as much as 1 msec.; after repeated A-stimuli the depression lasts still longer.

Internuncial Paths. The response to synaptic excitation is often complicated by the existence of alternative paths through interneurones, and as it is probably to such circuits that some of the characters of conduction through reflex arcs are due, we must briefly consider them. Elementary reflex arcs such as those depicted in Figs. 125 and 126 are merely convenient abstractions, and are probably not often operative, except in the myotatic reflexes. The common flexor, extensor thrust, and crossed extension reflexes all seem to be effected by multineurone arcs. Many neurologists believe, with Ranson and Hinsey,* that in addition to neurone relays intercalated on a direct path there are often multiple bypath interneurones (*principle of plurality of connections*) and also reverberatory side arcs or " re-exciting circuits " (*principle of reciprocity of connections*).† These important types of connections are illustrated in Figs. 147 and 148.

It is clear that circuits of these two types introduce new possibilities into the phenomena of conduction in the central nervous system, because, viewed in the simpler way, each synapse causes a delay of about 0·6 msec. As a result of the principle of multiple connections, an afferent impulse, in addition to being conducted direct to a terminal point, may also be delivered by detour paths to the same or other points and at later times, so that the excitation is spread out both spatially and in time (see

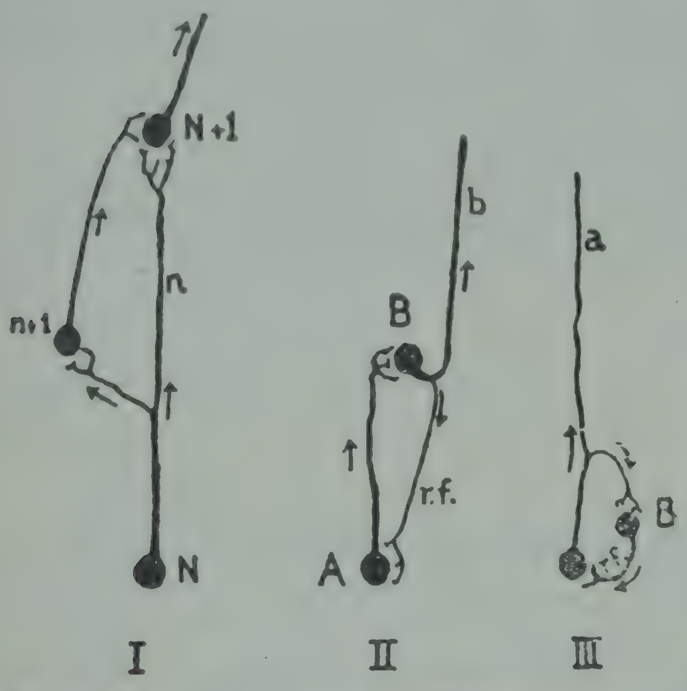


Fig. 148. Diagrams illustrating connections within chains of neurones. I illustrates " law of plurality of connections," II and III the " law of reciprocity of connections." Arrows show directions of impulses. (Lorente de Nó. *Arch. Neurol. Psych.*, 1933, **30**, 245.)

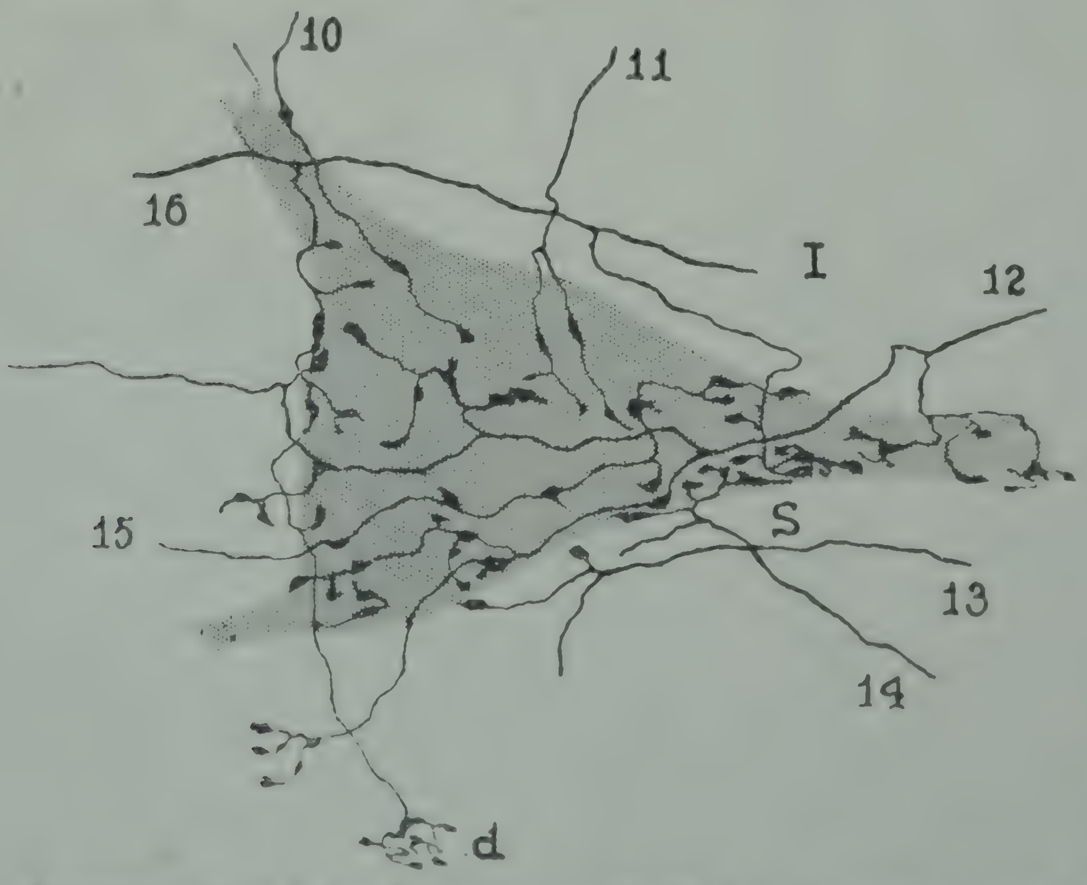


Fig. 149. Synapses on a large Internuncial Cell of the Spinal Cord of a Kitten. (Lorente de Nó. *J. Neurophysiol.*, 1938, **1**, 200.)

---

* Ranson and Hinsey. *Amer. J. Physiol.*, 1930, **94**, 471.

† Lorente de Nó. *Arch. Neurol. Psych.*, 1933, **30**, 245.

Fig. 151). In consequence of the closed chain, one impulse, or an afferent train of only a few impulses, can be worked up by re-excitation through delaying circuits into a longer train of impulses. With both types of circuit it should be noted that there results a slowly rising and more prolonged repetitive discharge of impulses into one or several points of a neurone pool.

It follows from this that the usual delay at a neurone pool, as measured by the time between arrival of the first presynaptic impulse and the resultant postsynaptic action, is likely to be much longer than the synaptic delay as evoked under the simplest conditions by a presynaptic volley. It is distinguished by the term *nuclear delay* (Lloyd).

Facilitation. The neurones of the oculomotor nuclei are most responsive, and their synaptic delay is shortest, when their c.e.s. is kept up by continual bombardment from internuncial neurones, too weak to actually excite the cells to discharge. Such cells are in a state of what is called by Sherrington "*subliminal excitation.*" Each cell in this condition is "on its toes," and ready, on receipt of a small additional stimulus, to discharge an impulse. This additional stimulus cannot come in the form of a second impulse at the same synapse, because the refractory period of the fibre is longer than the period of latent addition in the cell (see below), *i.e. temporal facilitation cannot occur* like it does in sympathetic ganglia.

Now if we consider the arrangement of the synaptic knobs in relation to the soma of a nerve cell, as shown in Fig. 149, it will be seen that a neurone receives contributions from several axon terminations, and that the synaptic knobs (of which there may be 1,000 or more on a single soma), are not necessarily terminal, but are distributed along the course of the fibrils in interlacing clusters of knobs surrounding the soma and dendrites, the synaptic terminals of different fibres often being at a distance from one another.

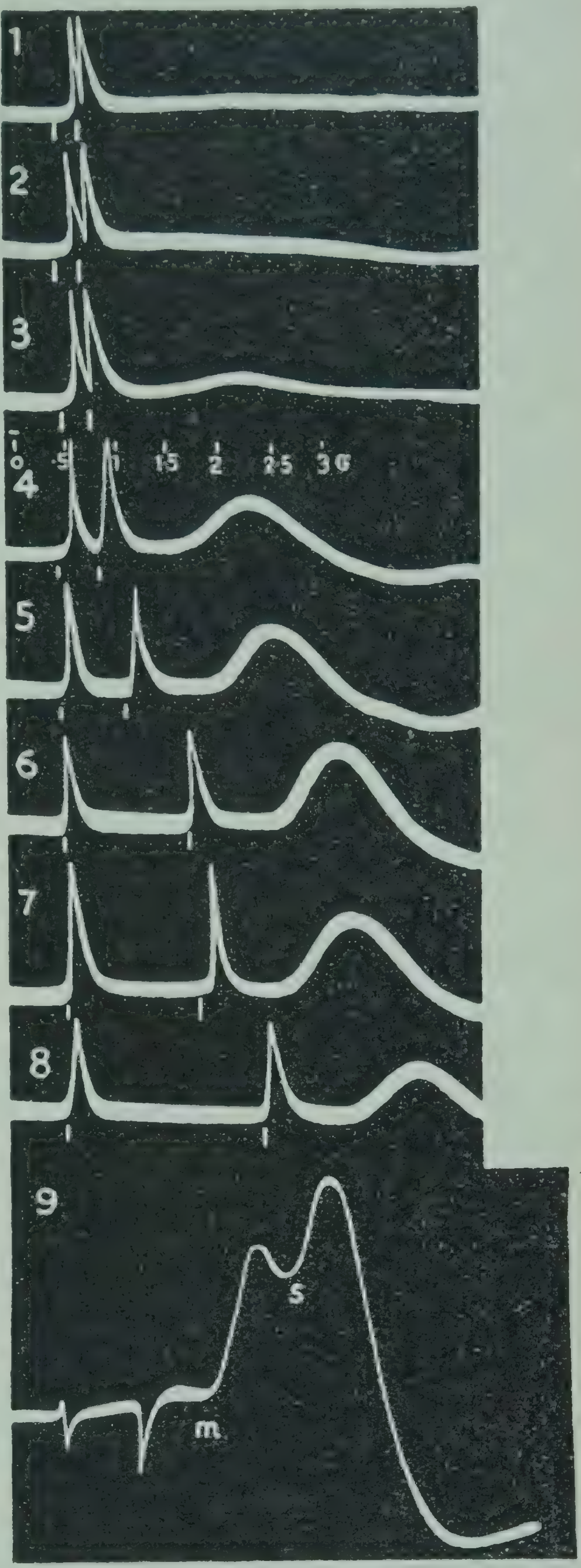


Fig. 150. Facilitation in motoneurones. *Curves* 1—8 : An F-shock, followed at the interval shown, by a C-shock. No effect at intervals below 0·25 msec., maximal effect at 1·25 msec. *Curve* 9 : Two C-shocks, the first one sub-threshold, the second one adequate, 1 msec. later. The m-wave the direct motor response, the s-wave the delayed synaptic response. (LORENTE DE NÓ. *J. Cell. Comp. Physiol.*, 1935, **7**, 47.)

It would seem probable, then, that facilitation would be seen when the later impulses arrived at new loci on the same soma. Evidence was, in fact, obtained by Lorente de Nó that usually more than one afferent fibre needs to be stimulated in order to activate, in any given zone of the soma, a sufficient number of synaptic knobs to cause excitation culminating in the discharge of a nerve impulse. A similar conclusion was reached by Sherrington for the motoneurones of the spinal cord. If too limited a zone at the surface of the neurone is excited the excitation remains a local non-propagating one which presumably dies down unless reinforced within a brief period (0·2 msec.), just as occurs with the non-propagated disturbances in nerve fibres.

In consequence of this it often happens that the stimulation of motoneurones by one set of synapses may be too weak and localised to result in excitation. Many of the motoneurones of the oculomotor nucleus are normally in a state of sub-liminal excitation as a result of the receipt of impulses from internuncial neurones, *e.g.* those from the labyrinth, and under these circumstances a stimulus applied at F in Fig. 146 is able to excite them.

This "tonic" subliminal excitation is abolished by destruction of the reticular substance of the medulla and pons, in which the internuncial paths lie. When this has been done a single shock at F gives no response, though two shocks 1 to 1·5 msec. apart usually do. Similarly, if a subliminal electrical stimulus is given to the motoneurones at the point C_1 (Fig. 146), and followed by a subliminal stimulus *viâ* the synapse by stimulating at F, say, 1 msec. later, a response may ensue.

At first sight the fact that facilitation occurs best when F-shocks are as much as 1 msec. to 1·5 msec. apart seems to conflict with the statement made above that the maximum period of latent addition of subliminal stimuli is 0·2 to 0·3 msec., after which the local c.e.s. dies away. The explanation shows that the matter is less simple than appears at first sight, and is as follows: The F-shock causes local synaptic excitation, and this dies away in 0·2 msec.; but it also excites a by-pass neurone having its own synapse, *viâ* which a delayed impulse reaches the motoneurone about 0·6 msec. later; if there are two by-pass neurones in series the delay will be at least 1·2 msec., and if at this moment, or within a short time after it, a second impulse arrives by the direct fibre from F, latent addition will occur. Facilitation, according to this, is therefore really due to the arrival, by two different paths, of almost simultaneous impulses at different synaptic points on the same soma. If in addition to these by-pass circuits we have closed reverberatory circuits the bombardment of the motoneurones can be kept up for much longer periods, so that facilitation may still be possible 30 msec. or more after a first stimulus.

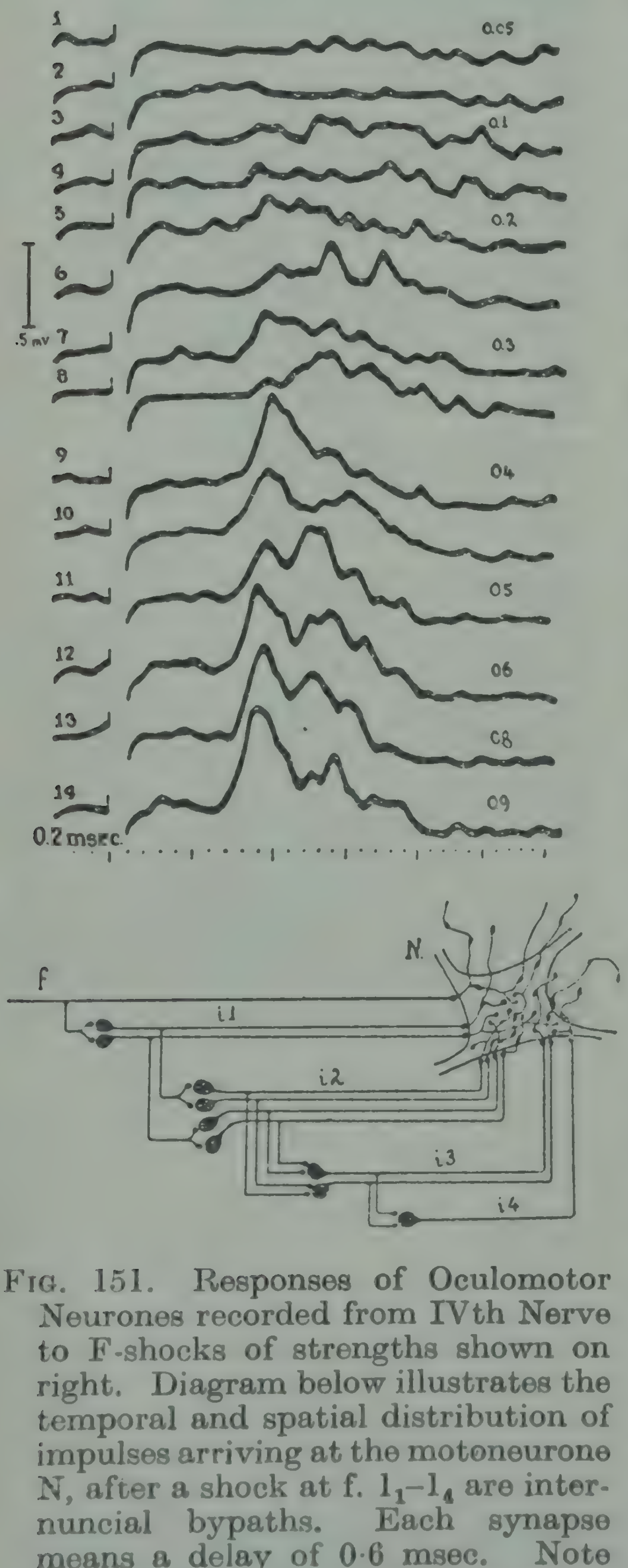


FIG. 151. Responses of Oculomotor Neurones recorded from IVth Nerve to F-shocks of strengths shown on right. Diagram below illustrates the temporal and spatial distribution of impulses arriving at the motoneurone N, after a shock at f. 1_1–1_4 are internuncial bypaths. Each synapse means a delay of 0·6 msec. Note that each fibre has several synaptic knobs on the soma. (LORENTE DE NÓ. *J. Neurophysiol.*, 1938, **1**, 207.)

This phenomenon of facilitation is well shown by Fig. 150, in which a subliminal F-shock is followed at varying intervals by a subliminal C-shock, *i.e.*, by a direct shock to the motoneurones. It will be seen that the earliest indication of a response was when the interval was something over 0·25 msec. Up till then the time is occupied by transmission from the F electrodes, say, 0·1 msec., followed by a lowering of the threshold of the motoneurone which reaches its maximum after about another 0·15 msec. After this we may suppose that delayed-path impulses continue to arrive at the motoneurones, and reach a maximum at 1·25 msec., afterwards declining. The effect of the C-shock is therefore greatest when it follows the F-shock at about 1·25 msec. (curve 7, Fig. 150).

It often happens that the latent period of the response of a motoneurone is considerably reduced by a previous excitation. This is not because facilitation shortens the synaptic delay, which is almost unaltered; it is because facilitation enables excitatory impulses to be effective directly without further (delayed) reinforcement through internuncial neurones.

It would seem that *the multiple chain of neurones is a normal unit of transmis-*

sion in the nervous system, and supersedes the classical and diagrammatic reflex arc of a fixed number of synapses. If the number of neurones in use in the chain at any time is large, facilitation may be continuous, and any applied stimulus may be followed by an "*after-discharge*." As is shown in Fig. 151, in which there was a background of internuncial bombardment, an increase in the strength of an F stimulus leads to a more sustained response, with shorter latent period. This response, however, even with the strongest stimuli, never engaged more than about a third of the possible motoneurones. As the Oxford school would say, the motoneurone pool response was *fractionated* into an active and an inactive group, the latter, or part of it, representing a *subliminal fringe*.

Inhibition. Sometimes, after a strong conditioning F-shock, there is a brief period of facilitation followed by a period in which responses to the testing shock are reduced or abolished. The explanation* of this phenomenon is that a refractory state and a state of subnormal excitability, as we have seen, follows the excitation of a neurone. Hence, since the first stimulus was a strong one, many internuncial neurones became involved—hence the early facilitation. But after this those neurones became refractory, and so the second shock from F received no facilitation and became ineffective. *Inhibition, therefore, is in this case not an active process happening in the neurones, but is merely lack of facilitation.*

In some parts of the central nervous system inhibition is very long lasting. Thus in the cerebral cortex a conditioning shock leads to prolonged extinction of the response to a testing shock.†

Transmission in the Spinal Cord. The phenomena of transmission in the spinal cord are necessarily more complex than those we have so far considered, because there are many possible paths.

When a single volley of impulses is fired along an afferent nerve, *e.g.* by giving a weak single shock to the central end of the popliteal nerve in a spinal cat, a reflex twitch of the ipsilateral flexor muscles, *e.g.* tibialis anticus, results. If a record of the electrical potential changes in the muscle (or in the motor nerve) is made, it is found that, if the afferent shock was weak, one efferent volley only is caused, *i.e.* the efferent fibres, or some of them, have each conducted one impulse in response to one afferent impulse.‡

If a very strong stimulus is given to the afferent nerve, then two impulses may pass along it, two also emerge by the motor path, and the muscular contraction is longer because of summation. But it also happens that a strong stimulus may cause only one afferent impulse in each fibre, but two or more efferent ones, with a resultant longer muscular response. This must mean that in some manner the motoneurones have been caused to fire at more than one moment in response to one afferent volley—a result called *repetitive discharge.* This phenomenon is readily explicable in the light of the information given in previous pages and does not necessarily mean that all the individual neurones are firing repetitively. The stronger stimulus excites more afferent fibres and causes impulses not only along the more direct paths but also along new internuncial paths, by which a belated volley of impulses arrives at the motoneurone pool, and excites the same or a fresh set of motoneurones, so that a later volley of impulses is fired either along the same efferent fibres, or along new ones not previously employed.

When two weak stimuli are given in succession to an afferent nerve for the spinal flexor arc, two impulses emerge from the efferent side if the two stimuli are more than 1·0 to 1·9 msec. apart.§ If the interval is less than this, the second stimulus is ineffective because it falls in the refractory period of some part of the reflex arc. Actually this refractory period (absolute and relative) may be as much as 4 to 6 msec., but when two or more afferent stimuli are given at, say, 1 msec. intervals, the second one travelling in the wake of the first is conducted more slowly, so that it only arrives at the centre about 6 msec. after the first, *i.e.* not until the arc has recovered.

---

* Lorente de Nó. Symposium dedicated to Professor Beritoff, Academy of Sciences of U.S.S.R. Georgian Branch, Tiflis, 1936, 231.

† Dusser de Barenne and McCulloch. *Proc. Soc. exp. Biol. Med.*, 1934, **32**, 529.

‡ Adrian and Forbes. *J. Physiol.*, 1922, **56**, 301.

§ Adrian and Olmsted. *J. Physiol.*, 1922, **56**, 426.

It cannot be said precisely where in the arc the refractory state is longest. There is, of course, a refractory state in the nerve, but the absolute refractory period is only of the order of 0·4 msec. and the relative about 3 msec. for the fast fibres. It would seem probable that the neurones of the spinal cord have very similar properties, and that owing to the delayed conductance for second impulses these do not reach the neurones until their refractory period has passed off.

If a series of stimuli are given to the afferent nerve, so that a train of impulses passes along it, the efferent nerve also conveys a train of impulses at the same frequency, provided the afferent stimuli are not too close together. The minimum interval between stimuli in a series is about 3·5 msec., *i.e.* the maximum frequency at which a series of afferent nerve stimuli can be translated into efferent ones is about 300 per second.* It would appear then that if, in these experiments, the same motoneurones respond to each volley, these motoneurones cannot be made to send out a regular series of impulses at a greater frequency than about 300/sec. (and if the spinal cord is cooled the rate is slower). The muscle, as we already know, is able to respond to stimuli faster than this, so that there is no question of the motoneurones being able to discharge impulses more rapidly than the muscles can follow them.

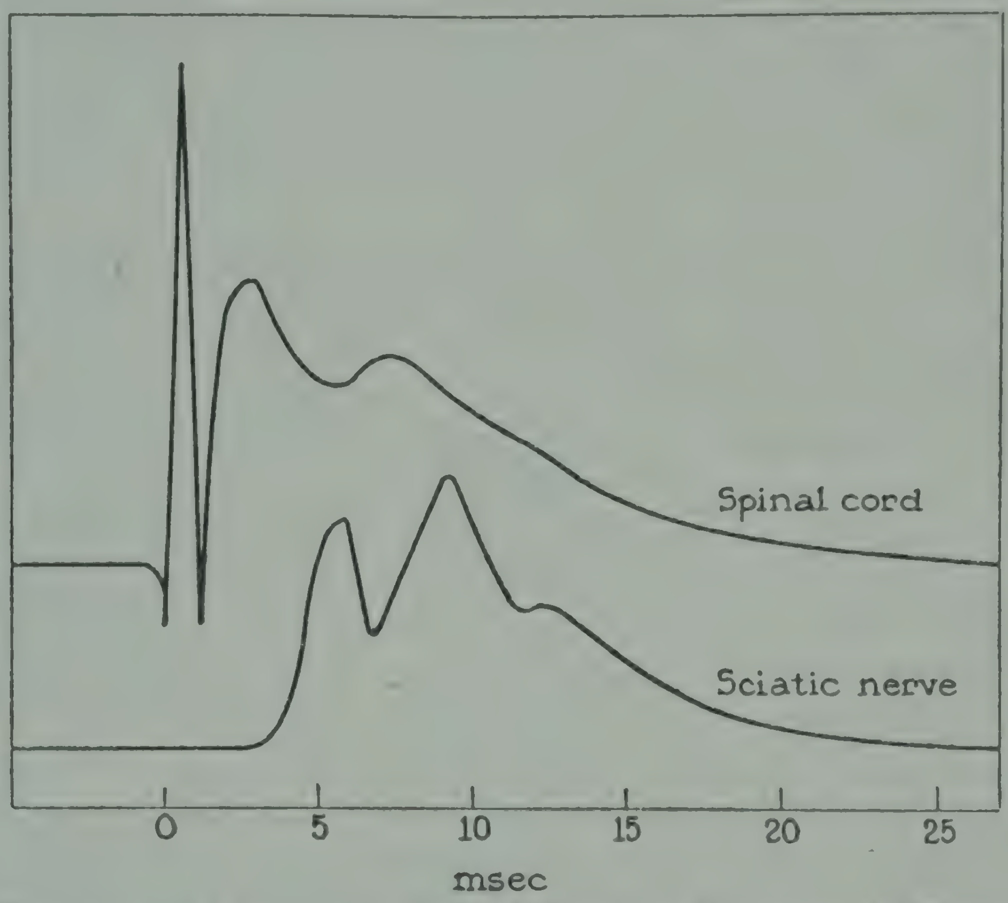


Fig. 152. Relation of Afferent Spike and Negative Internuncial Potentials in the Lumbar Cord of the Spinal Cat to the Reflex Discharge into the Sciatic Nerve. The positive after-potentials come later and are not shown. Gasser. Harvey Lectures, 1936, 182.

In reflex contractions brought about by ordinary sensory stimuli, *e.g.* pinching the foot, the impulses pass along the motor nerves at a frequency up to 100 to 250/sec. With frequencies of this order the muscle, of course, gives a perfectly smooth contraction, and the component excitations can only be seen in the electrical record.

Cord Potentials. Studies of the electrical potentials of the cord yield further important information. If an afferent volley is sent into a dorsal root, it may be traced as a spike potential into the cord and for some distance along the dorsal columns. Leads from the spinal cord surface show that this spike potential is followed by slow waves, which according to Gasser are due to the excitation of internuncial neurones.* The potentials produced in the motoneurones are obscured by these, but their discharge is indicated by the volley which finally emerges from the ventral root. The original afferent spike, the discharge potentials of the internuncials, and the efferent discharge along the motor nerve are shown in Fig. 152. It will be seen that the discharge along the efferent nerve runs parallel with the curve of negative cord potential. This negative phase is attributed to the excitation of internuncials and not of motoneurones, because an antidromic impulse which would discharge the latter, does not modify it. During the time that there is a negative potential (10 to 20 msec.) the internuncial neurones are emitting impulses which pass on to the motoneurones. According to Hughes and Gasser † the negative cord potential represents the overlapping of these internuncial spikes. During this period the reflex arc is facilitated. Each spike discharge of an internuncial is presumably accompanied by a refractory period, after which there is no demonstrable negative after-potential comparable with that of fresh nerve, and accordingly there is no supernormal phase like there is in fresh nerve fibre, but instead, there is an immediate subnormal state, accompanied by prolonged (100 msec. or more) positive after-potential (Fig. 153), as is seen in nerve fibres after many previous stimuli (*v.* p. 177, Fig. 97). During this period the reflex arc is therefore in a state of depressed excitability, so that a second volley, say 50 msec. after the first, causes a smaller response, *i.e.* the reflex is *inhibited*.

---

* Cooper and Adrian. *J. Physiol.*, 1924, **59**, 61.

It will be seen later that these internuncial waves negative (10 to 20 msec.) and positive (70 to 200 msec.) correspond roughly with the time course of c.e.s. and c.i.s. respectively.

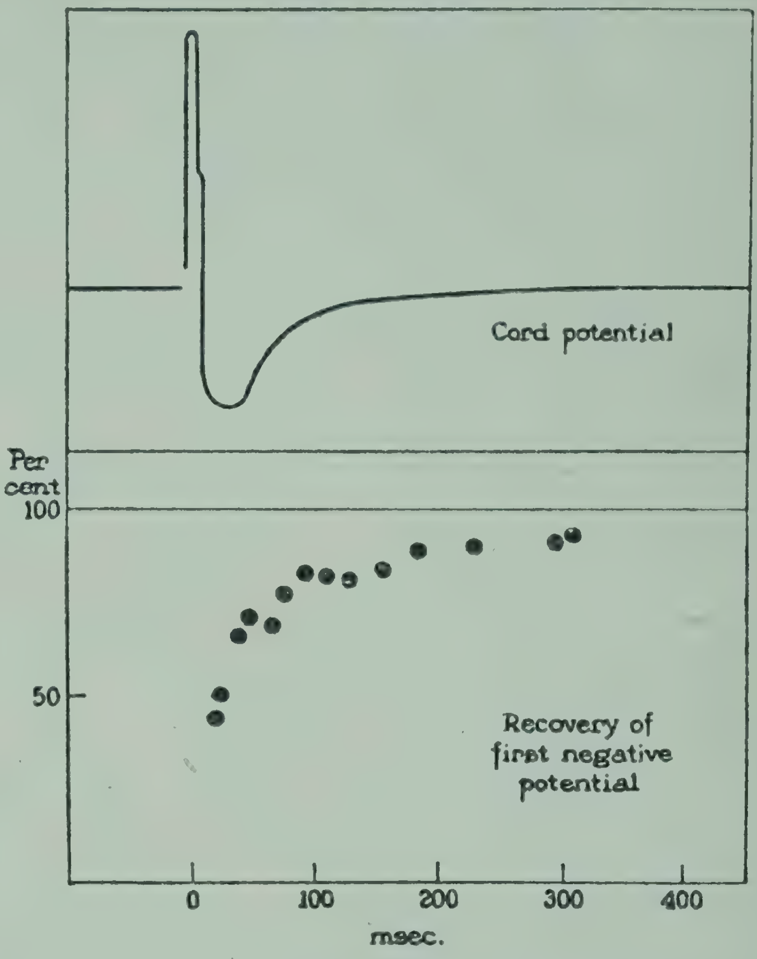


FIG. 153. *Above*, Course of the Cord Potential in response to one Volley. *Below*, size of negative internuncial response caused by a second volley to same root. Internuncial negative potential is reduced during period of positive after-potential. (HUGHES and GASSER.)

Root Potentials. A somewhat different interpretation of spinal cord potentials is given by Barron and Matthews,‡ who recorded the potential changes in the dorsal and ventral nerve roots, when afferent impulses entered the cord as a result of either electrical stimuli to the dorsal roots or natural stimuli to the nerve terminals. In both cases they found that (apart from spike potentials), slow negative potential waves occurred, *i.e.*, the electrode nearer the cord was negative to the other. The waves last for from 0·1 sec. up to periods of seconds, and are seen in both dorsal and ventral roots whenever these are involved in reflex actions (Fig. 154). The waves in the dorsal roots are seen not only in stimulated roots, but also in neighbouring roots on the same or the opposite side. Hence they cannot be caused by the passage of nerve impulses along all the nerve fibres that show the waves. The spike potentials, when they occur, of course, represent the passage of nervous impulses, and need not be further described. They are always superimposed on the slow negative waves, and these are larger and longer the more intense and prolonged the afferent stimulus. With very brief and weak stimuli the slow potentials may occur in the dorsal roots without any spikes appearing at all. What they indicate is a tendency to discharge spikes. These slow negative potentials, which are probably similar to those ascribed by Hughes and Gasser to overlapping internuncial spikes, are differently interpreted by Barron and Matthews. They regard them as being due to a depolarisation of the axons of the roots due to electrotonic spread from sites where an active depolarisation obtains. The roots merely act as leads by which electrotonic currents can be tapped off from these sites (*v.* p. 191). This state of depolarisation is an essential accompaniment of excitation, and in fact is the ultimate cause of it. When a motoneurone is sufficiently depolarised it emits nerve impulses rhythmically at a frequency depending, among other things, on the degree of the depolarisation. In evidence of this they adduced the fact that it is possible by the passage of a *constant* current to the surface of the cord (Fig. 155) to depolarise the motoneurones and so to cause them to discharge rhythmically (Fig. 156). In this respect the motoneurones are therefore different from normal nerve fibres and more closely resemble the sensory terminals.

* GASSER and GRAHAM. *Amer. J. Physiol.*, 1933, **103**, 303.
† HUGHES and GASSER. *Amer. J. Physiol.*, 1934, **108**, 295, 307.
‡ BARRON and MATTHEWS. *J. Physiol.*, 1938, **92**, 276.

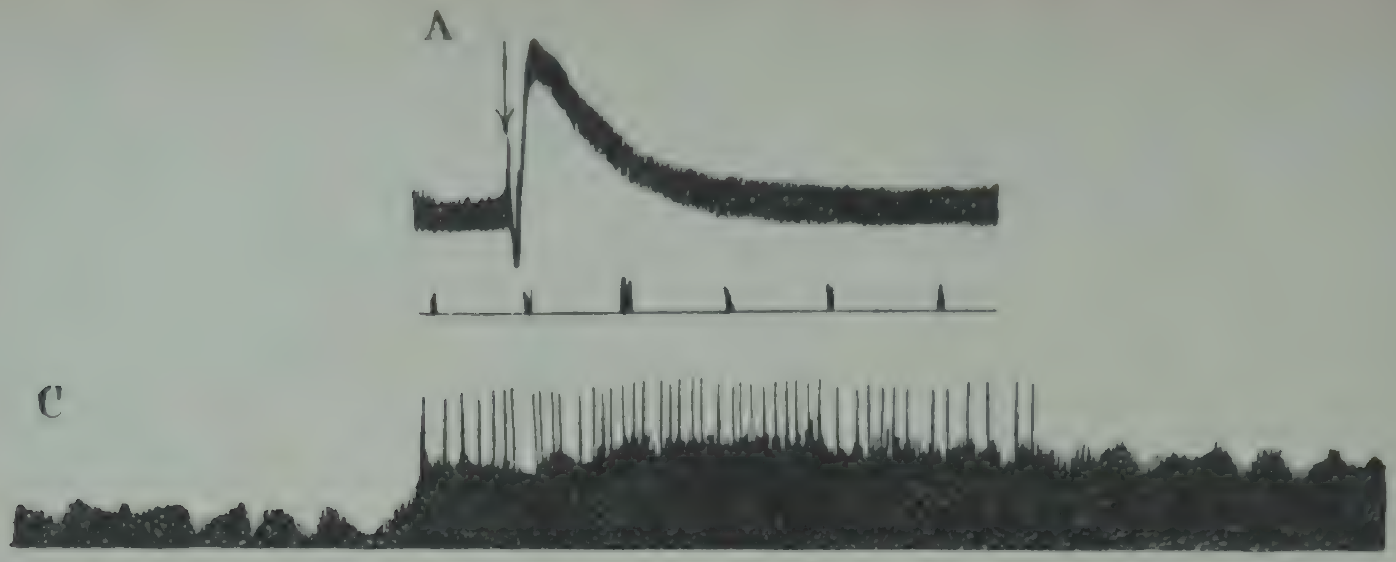


FIG. 154. A. Potential Changes in Dorsal Rootlet of Frog when Stimulated by Single Shock. Spike followed by slow potential. (Time = 0·1 and 0·02 sec. intervals.) C. Potential changes in ventral root of L_6 in cat (fibres reduced by partial section). Stimulus by squeezing toes. The shower of spikes lasts 0·7 sec. and is superposed on a negative potential wave lasting beyond it. (BARRON and MATTHEWS. *J. Physiol.*, 1938, **92**, 276.)

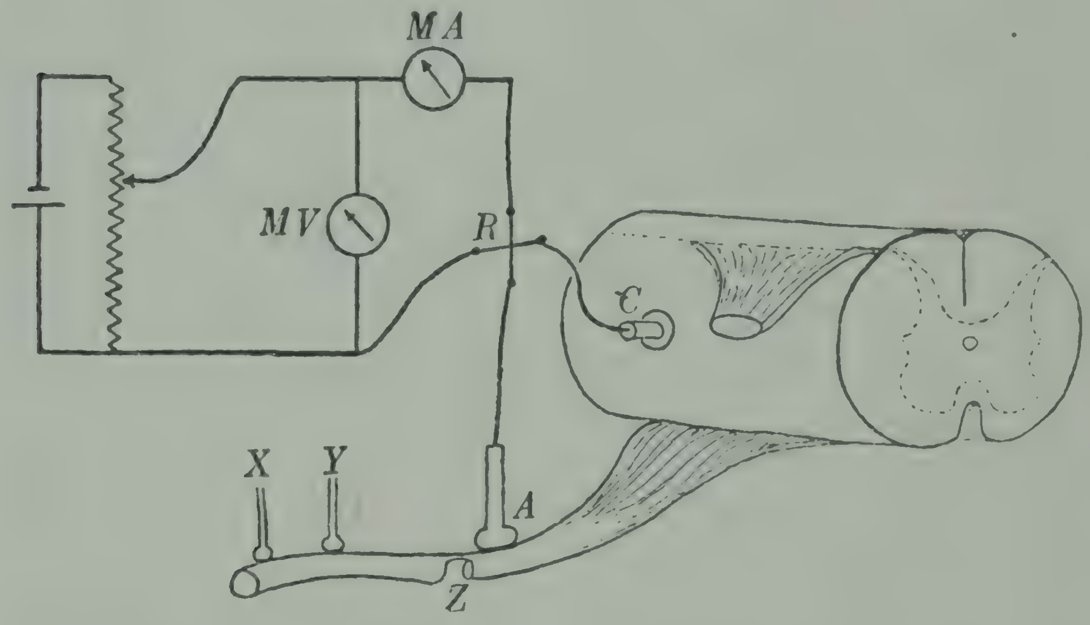


FIG. 155. Depolarisation of Motoneurones by Constant Current. *A*, anode; *C*, cathode; *X*, *Y*, recording electrodes; *Z*, point at which all active fibres but one have been divided. (BARRON and MATTHEWS. *J. Physiol.*, 1938, **92**.)

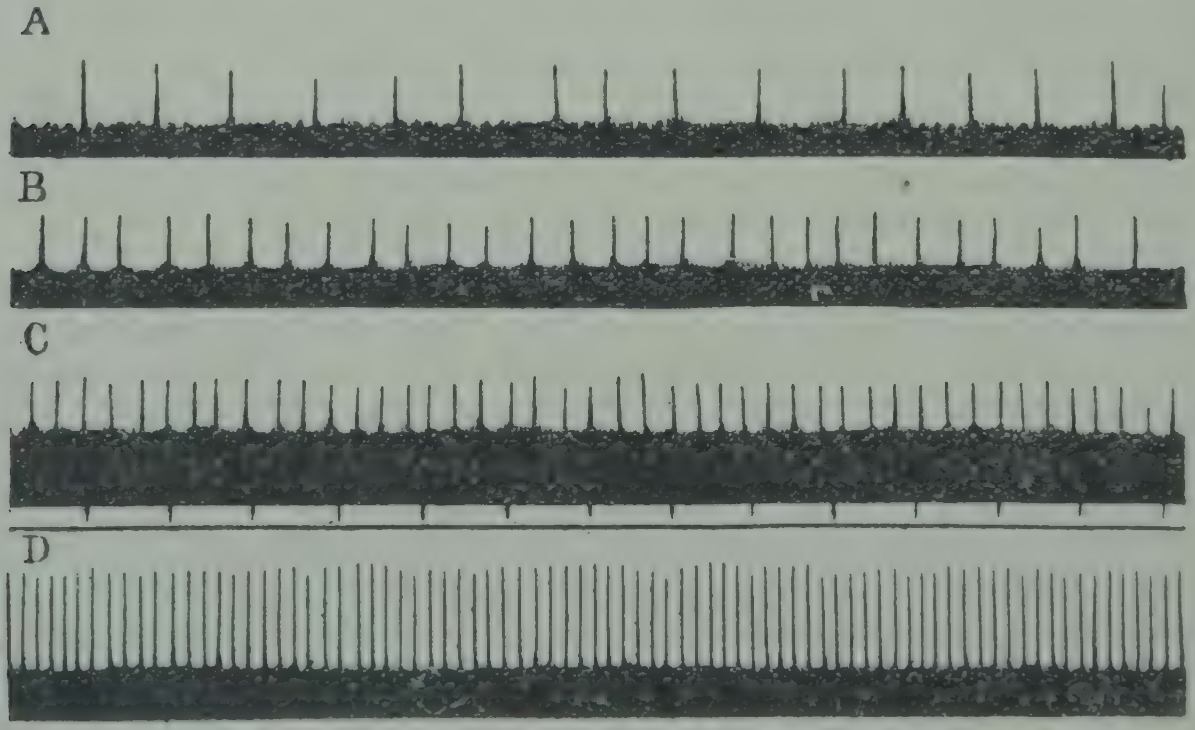


FIG. 156. Stimulation of Single Motoneurone of Frog by Constant Current of Strength increasing from A to D. (BARRON and MATTHEWS. *J. Physiol.*, 1938, **92**.)

Owing to electrotonic spread, when the ventral root is led off, the record indicates not merely the discharge of impulses, but, for a short distance along the root, also the depolarisation of the motoneurones which causes it. When the dorsal root is tapped the depolarisa-

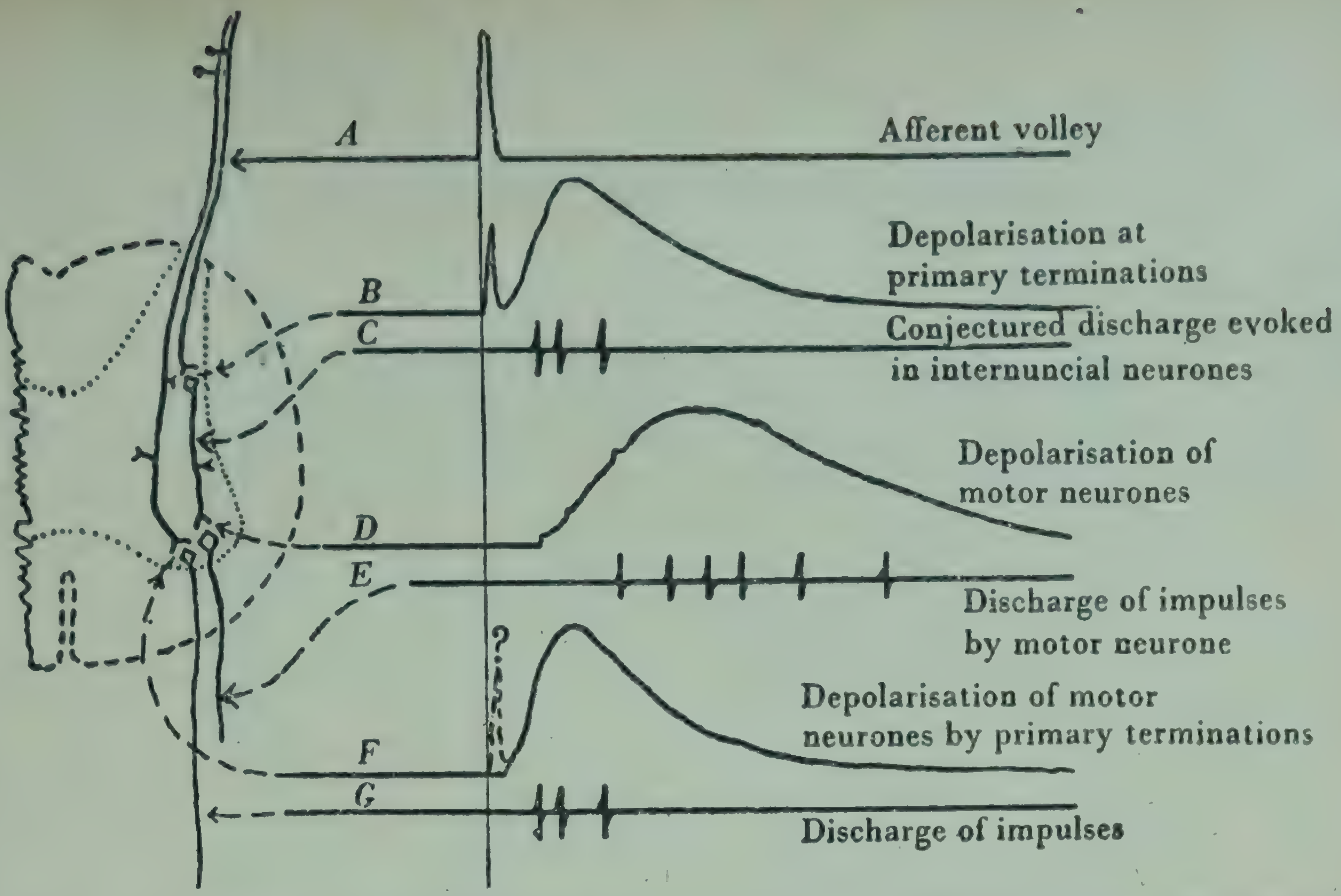

A
Afferent volley
B
Depolarisation at primary terminations
C
Conjectured discharge evoked in internuncial neurones
Depolarisation of motor neurones
D
E
Discharge of impulses by motor neurone
?
Depolarisation of motor neurones by primary terminations
F
G
Discharge of impulses

FIG. 157. Diagram summarising the Observations and Deductions of Barron and Matthews. (BARRON and MATTHEWS. *J. Physiol.*, 1938, **92**, 316.)

tion at the fine non-medullated feltwork in which the afferent fibres end also spreads electrotonically for a short distance along the root. The relation of these depolarisations to the transmission of impulses is indicated in Fig. 157.

For Reference

BODIAN (1942). Cytological Aspects of Synaptic Function. *Physiol. Rev.*, **22**, 146.
GASSER *et al.* (1939). Symposium on the Synapse. (From *J. Neurophysiol.*, 1939, **2**, 361. (Baltimore : Thomas.)
LLOYD (1944). Functional Organization of the Spinal Cord. *Physiol. Rev.*, **24**, 1.

CHAPTER XVII

CHARACTERISTICS OF REFLEX ACTIONS

In his classical "Integrative Action of the Nervous System," Sherrington showed in 1906 how reflex actions are organised and modified to serve the needs of the individual, and the continued researches of his school from that time to the present have further widened and deepened our knowledge of the phenomena of nervous action. Most investigations on reflex action have been on motor reflexes, because of the ease with which they can be studied quantitatively. A convenient measure of a response in a given muscle is the isometric tension produced when it contracts. We may consider a given muscle to be innervated from a group of ventral horn nerve cells, which is called the *motoneurone pool*.

The Motor unit. Each motor neurone has an axon which branches freely near its termination and so supplies many muscle fibres. These all contract together. This is an example of the principle of divergence. The actual number can be determined by dividing the number of muscle fibres in a given muscle by the number of nerve fibres in its motor nerve. The average is around 150 fibres per motor neurone, but is greater with powerful muscles (*e.g.* 500 in cat's gastrocnemius) and less with delicate and precise ones such as the eye muscles. A motor neurone with all its muscle fibres is called a *motor unit* or motounit. If the motor nerve to a muscle is given a maximal tetanic stimulation for a second or so, the tension produced when divided by the number of motor units, gives us the maximal tension per unit. This for the cat varies from 6 g. for semitendinosus to 30 g. for gastrocnemius; for simple twitches tensions about one-fourth as great result. A muscle like the cat's soleus having about 250 motor units and giving a maximal tension of about 2,500 g. can thus have 250 gradations of 10 g. each as its tension rises from zero to maximum.

CHARACTERISTICS OF REFLEXES. Let us first take as an example the flexor reflex already mentioned (p. 230); it illustrates certain properties common to them all.* We shall mention later the respects in which the extensor reflexes show points of difference. Stimulation of almost any afferent nerve in the same limb will evoke the flexor reflex. Thus, stimulation of the tibial (popliteal) nerve is commonly used to produce reflex contraction of the tibialis anterior, which is a favourite flexor for experiment: its motor nerve is peroneus.

(1) Localisation. The flexor-reflex is normally caused by a nociceptive stimulus applied to the same limb; the response is quite regular and automatic in appearance. It is determined by the fact that the nerve fibres, mostly of small or medium diameter, which carry the impulses from, say, the toe, into the spinal cord, divide there and make connections with the motoneurone pools whose axons are distributed to the several muscles involved in the reaction. The connection of the sensory with the motor neurone is not direct, since the impulse has to pass through at least one intermediate neurone before arriving at the motor neurones. The main path of the impulse, however, is definite and unvarying, despite the fact that innumerable other paths are theoretically possible, and that the reflex may spread into these if afferent stimuli are powerful.

(2) Co-ordination. In the flexor response all the physiological flexors are brought into action. A reflex achieves, not the contraction of a particular muscle, but the performance of a significant movement. This always involves *groups of synergistic muscles*, and is brought about through the agency of multi-neurone arcs.

The effect varies with the strength of the stimulus. The minimal effective stimulus causes simply flexion of the foot. As its strength is increased, this is attended by flexion of the leg on the thigh, and finally by flexion of the thigh on the body. With still stronger stimulus there is a spread to the opposite hind limb, which, however, extends. Increase in the strength of stimulus causes not only an increase in the strength of contraction of the reacting muscles, but also a spreading of the reaction to more and more muscles, or groups of muscles. The spread occurs always in definite order. The stimulus when represented by the prick of a needle can affect only one or two nerve fibres. The impulse carried along these spreads in the cord, reaches the motor neurones and causes them to discharge. The first discharge is, as a rule, limited to those in the immediate proximity of the entering impulses, but, even when minimal, involves the simultaneous action of more than one ventral root. We may say that the response is determined to a certain extent by the

---

* Eccles and Sherrington. *Proc. R.S.*, 1931, **107B**, 511, 535, 586, 597.

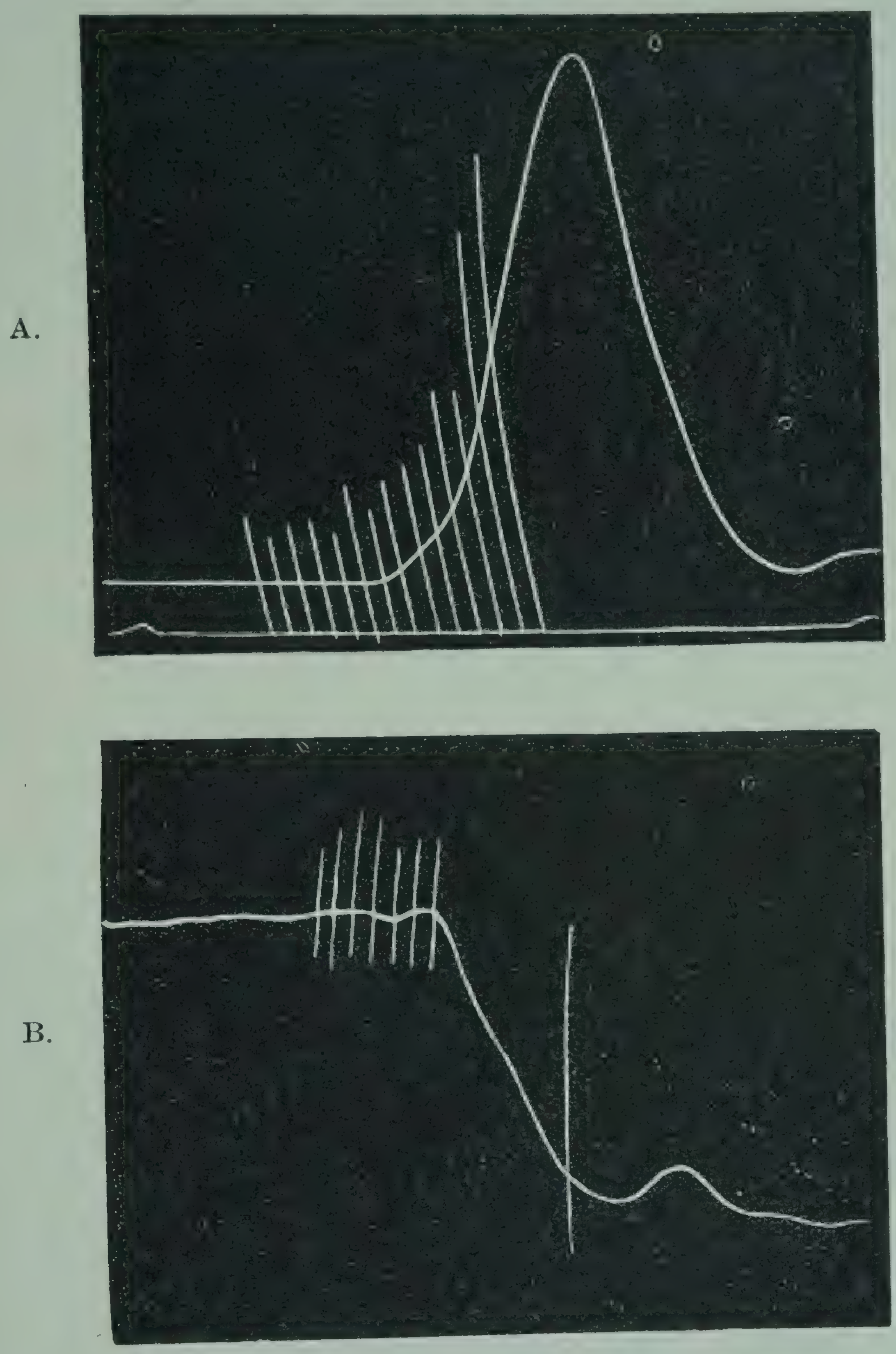


Fig. 158, A and B. (Sherrington.)

The Flexion Reflex observed as Reflex Contraction (Excitation) of the Flexor Muscles of the Knee (A), and as Reflex Relaxation (Inhibition) of the Extensor Muscle (B).

The stimulus was a series of weak break induction shocks applied to a twig of the internal saphenous nerve below the knee. Observation B was made four minutes after A. Note the summation of stimuli, in each case six stimuli being required before the reaction was evoked.

proximity of the afferent to the efferent paths, but that it is nearly always plurisegmental, the most important determining factor being the adaptation of the movement to the stimulus.

The gradual spread of the response with increasing strength of stimulus is spoken of as "*irradiation.*" The nature of the response is determined by the place of application of the stimulus, and by the quality of the latter. While a painful stimulus causes flexion of the leg, firm pressure on the plantar surface of the paw causes the extensor thrust—the "stepping" reflex. However extensive the irradiation, the muscles which are set into action are always such that their actions co-operate towards a given end. Thus, when the impulse spreads to the opposite limb and produces an extension, the reaction is such as would ensue when the animal steps on a sharp point and immediately retracts the irritated limb away from the injurious agent, while it extends the other limb as support, and in the first act of movement away from the spot.

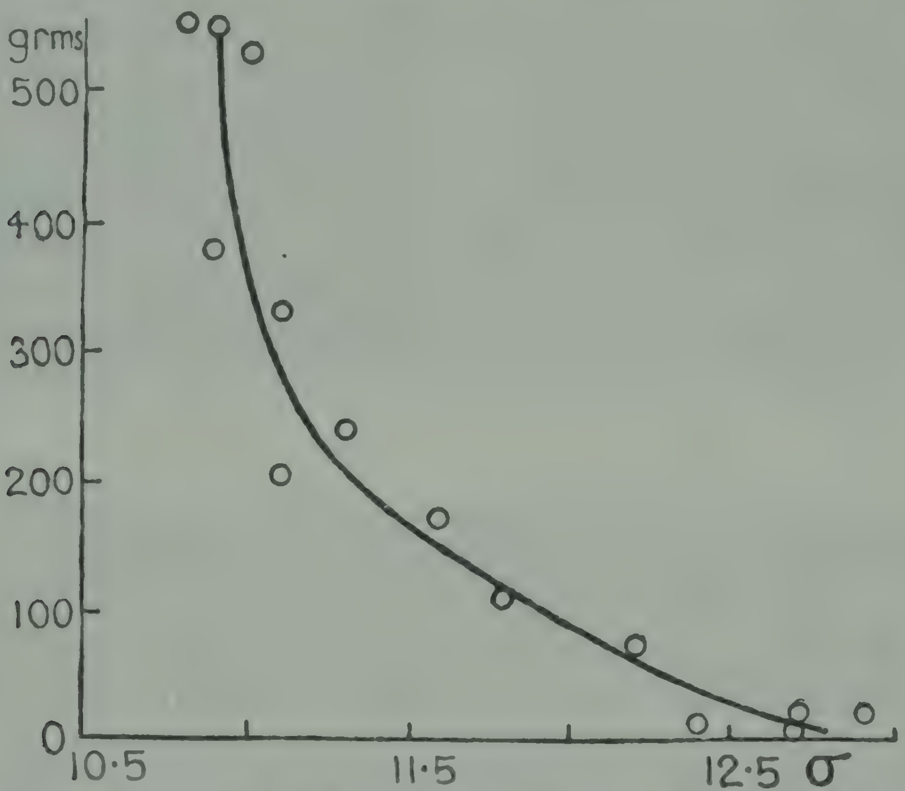


Fig. 159. Effect of Strength of Stimulus on Latent Period of Reflex Response. Reflex excitation of tibialis anticus through tibial nerve of same side. The strength of stimulus is indicated in the figure by the strength of the response. (Creed *et al.* "Reflex Activity of the Spinal Cord." Oxford Univ. Press, 1932.)

(3) **Reciprocal Innervation.** A superficial study of this reflex might lead us to suppose that, by the varying resistance in different synapses, the impulses are directed so as to affect exclusively the muscles whose contraction will co-operate in the primary reflex. Such a description, however, would represent only one-half of the process. Every muscle in the body is in a variable state of tone. If the tone in the opponent muscles is not to interfere with the carrying out of a reflex movement, there must be some means by which it can be inhibited. Such an inhibition always occurs, and the remarkable fact was brought out by Sherrington that the impulses, which start on the surface of the body and set loose the co-ordinated contraction of certain muscles, spread at the same time to the motor mechanisms governing the muscles antagonistic to the movement, and exercise on these an inhibitory effect.

This inhibition can be best shown in the following way: The extensor thigh muscles are cut away from their attachments to the tibia, and the patellar tendon is connected with a recording lever. The flexor muscles of the knee are similarly detached from their insertions and connected with another lever. On then exciting the flexor reflex by stimulation of an afferent nerve of the leg, the lever attached to the flexors shows a contraction, while that connected with the patellar tendon falls (Fig. 158), showing that the extensor muscles have undergone actual elongation. The inhibition of the extensor tone is thus a direct result of the primary stimulus.

(4) **Delay.** When the latent period that elapses between stimulus and response in the flexor reflex is measured, and the time taken in conduction along the nerve fibres is deducted, we get what is called the *central reflex time.* Thus if the latent period was 10·4 msec., and the nerve conduction time was 6·5 msec., then the time spent in central processes was 3·9 msec. This time is spent in two ways ; first, and chiefly, in raising the c.e.s. of the neurones concerned, and secondly in true synaptic delay, which is probably quite brief. We find that central reflex times vary greatly according to the particular reflex ; *i.e.* to the number and properties of the neurones, and also according to the strength (Fig. 159), number and frequency of the stimuli given. The flexor response is a brisk one, with central delay of 3 to 6 msec. Contralateral reflexes usually occupy a longer time, but in all cases the stronger and more frequent the stimuli the shorter the latent period will be. We may suppose that this is because with stronger stimuli a given motoneurone receives a greater number of impulses from different synaptic knobs in a given time.

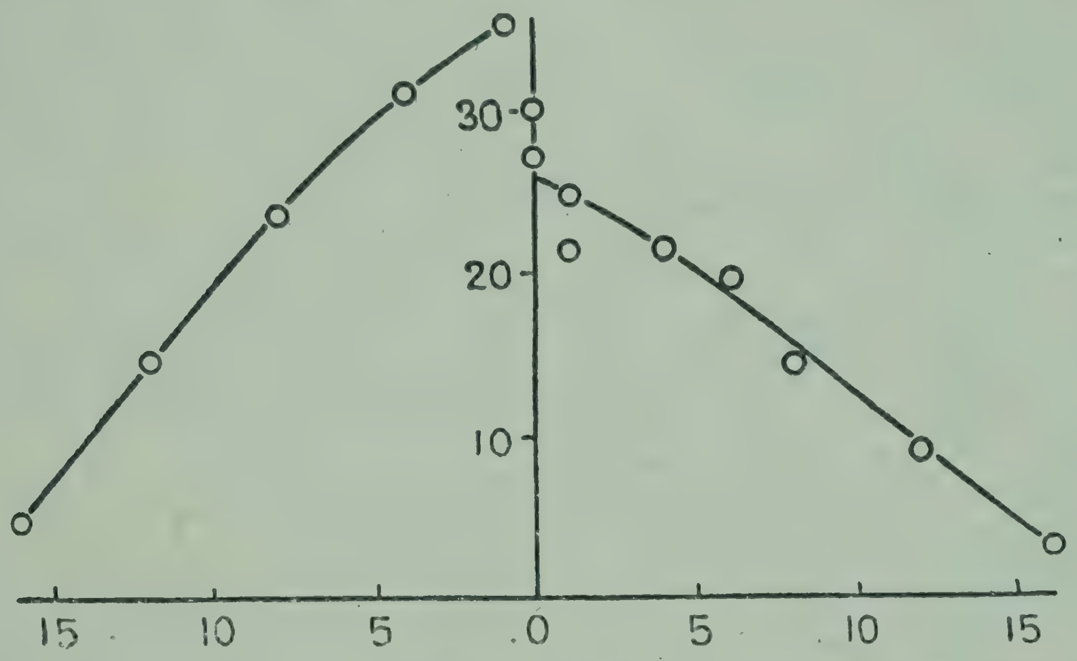


Fig. 160. Reflex Contraction of *Tibialis Anticus* to Pairs of Threshold Stimuli separated by the Intervals shown on the Abcissæ. Ordinates, tension in g. The stimuli were given first to medial gastrocnemius nerve, second to lateral gastrocnemius nerve (on left of zero) and in reverse order (to right of zero). (Creed *et al.* "Reflex Activity of the Spinal Cord." Oxford Univ. Press. 1932.)

When natural stimuli are given the afferent impulses are not volleys, but are dispersed, so that latent periods may be quite long.

(5) **Summation** of stimulation is a conspicuous feature of reflex actions. A single shock of a strength which would excite strong contraction of the tibialis anticus when applied to its motor nerve (the peroneal) is often found to be without effect when applied to an afferent nerve such as the tibial.

Summation may be obtained when two stimuli are applied to different afferent nerves, and experiments on summation are usually made in this way (Fig. 160), because then there can be no question of the nerve fibres being in a refractory state following the first stimulus. If the two stimuli are simultaneous, this type of summation is sometimes called *spatial summation* or "immediate induction." Its occurrence illustrates the fact that different afferent paths may converge to a final common path. The facts of summation are represented diagrammatically in Fig. 161.

It will be seen from this figure that when a fibre is excited, in addition to affecting certain neurones in the pool an area of subliminal c.e.s., called the *subliminal fringe*, is created around these. When a second fibre commanding the same reflex is stimulated at the same time, it also creates a subliminal fringe, and if and where these adjacent fringes overlap, the result is that the c.e.s. there will rise to a sufficient level to cause excitation.

Although, as we have seen, temporal summation at the synapse itself does not occur in the C.N.S., yet if the two (or more) stimuli, instead of being given simultaneously, are given in succession, summation (temporal summation, or "successive induction") can still occur if the interval between them does not exceed a certain amount, in the case of the flexor reflex about 15 msec. (Fig. 160). This may be interpreted as meaning that the c.e.s., which takes about 4 msec. to be established, will all have declined after 15 msec. In some reflexes (*e.g.* scratch reflex) Sherrington found that as many as forty stimuli may be needed to release a reflex, and the latent period may then amount to seconds. Similarly, two stimuli, each of which gives only a small response, may summate to give a larger one. Naturally the greatest effects of two stimuli are seen when they are simultaneous, because then the c.e.s. due to each reaches its maximum at the same time.

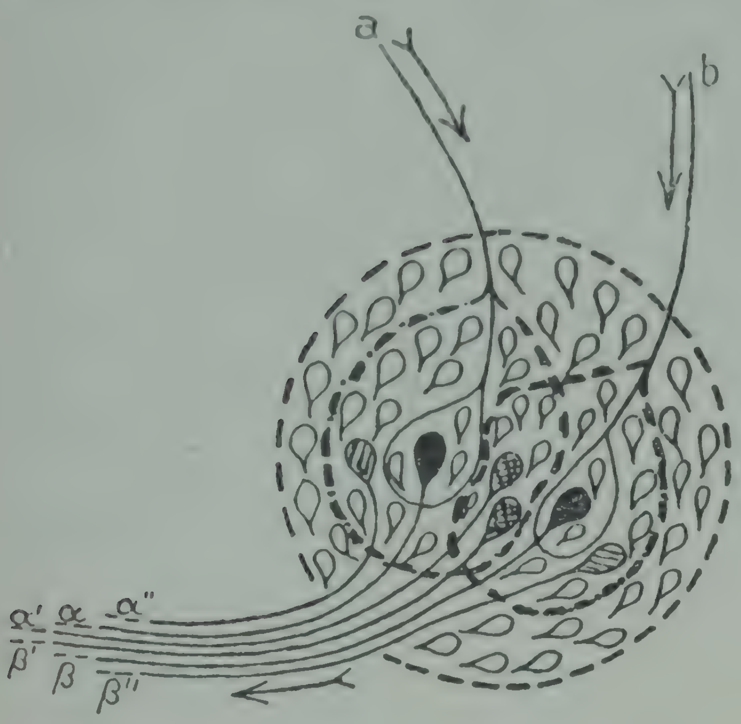


FIG. 161. Two Afferent Fibres, *a* and *b*, are shown entering a Motoneurone Pool, *α*, *β*, etc. are efferent fibres. The continuous line shows the field of threshold action (one neurone excited in each). The dotted lines inside the pool show the subliminal fields (three neurones in each). Each fibre alone excites only one neurone ; concurrently stimulated four neurones are excited, *i.e.* one from each threshold field and two from overlap of the subliminal fringes. (After CREED *et al.* "Reflex Activity of the Spinal Cord.")

Owing to the facilitating effect of the first volley, the central reflex time of the response to the second is shortened even to less than 0·5 msec.

It will be clear that the explanation of summation is to be found in the facilitation phenomena which we have already studied. The succeeding stimuli no doubt open up delaying by-paths through internuncial neurones until finally a sufficient number of impulses arrive within a given time at the motoneurones concerned.

Successive spinal induction is said to be positive when a previous stimulus favours a succeeding one, as in the examples we have mentioned. In some cases, however, where two reflexes are opposed, a reflex will favour the appearance of the opposing one. Thus a flexor reflex may be followed by an extensor movement, thus revealing a pre-existing state of subliminal extensor excitation. This is called *negative successive induction*, and such phenomena are probably responsible for rhythmic reflexes such as those of stepping.

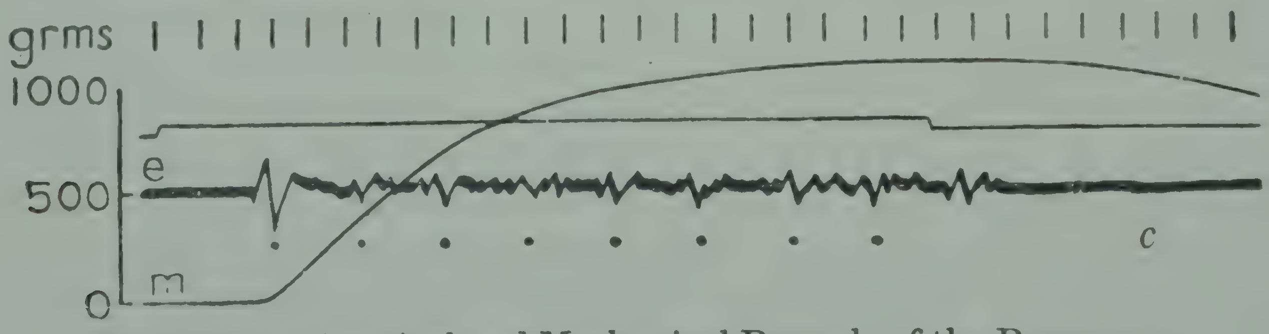


FIG. 162. Electrical and Mechanical Records of the Response of Tibialis Anticus to Repetitive Stimulation of Posterior Tibial Nerve. Frequency of stimuli marked by dots and large waves in electrogram ; note the secondary waves. Time 10 msec. (CREED *et al.* "Reflex Activity of the Spinal Cord." Oxford Univ. Press, 1932.)

(6) **After-discharge.** When the motor twitch of the tibialis anticus caused by stimulation of its motor nerve by a single shock is compared with the reflex contraction caused by a single powerful shock to the tibial nerve, it is seen that the reflex contraction lasts for about 0·1 sec. longer than the motor-nerve twitch. This is due to what is called "after-discharge" and in extensor reflexes is very much more prolonged still.

It is also longer after repeated stimulation than it is after a single afferent shock (Fig. 162), and electrical records show that there are efferent nerve impulses after stimulation of the afferent nerve has ceased.

The cause of after-discharge has already been discussed when dealing with the properties of central conduction. It is due to the fact that the impulses travelling to the motoneurones are delayed to varying extents.* Some arrive almost direct, others are delayed in internuncial by-paths and reach the motoneurones later (Fig. 163); if the refractory period of these has passed off, these later impulses may cause them to fire off another impulse. Hence there are more efferent impulses than afferent stimuli (Fig. 162); with strong stimuli some of the later impulses also reach motoneurones not previously affected, and so the motor discharge is prolonged.

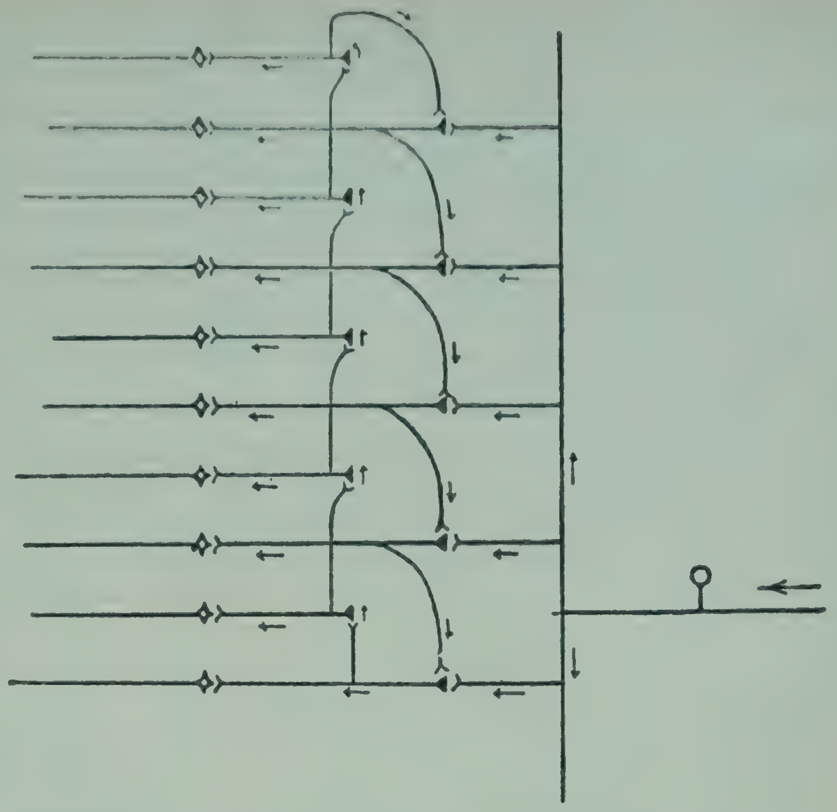


Fig. 163. Diagram to Illustrate Causation of After-discharge by Passage of Impulses along Delay and Re-exciting Circuits. Arrows show directions of impulses but not their time-relations; these latter depend on the number of synapses passed. (Ranson and Hinsey *Amer. J. Physiol.*, 1930, **94**, 471.)

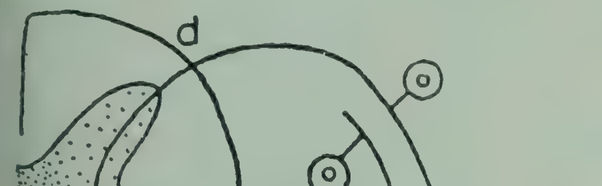


Fig. 164. Schema of Arrangement for Firing Antidromic Volley. v, d, ventral and dorsal roots; a, electrodes on motor nerve; c, electrodes on an afferent nerve. Needle electrodes stuck into muscle lead to a string galvanometer for recording action potentials. (Creed *et al.*)

If an antidromic volley is fired back from the motor nerve, the after-discharge lapses for from 20 to 60 msec. and then reappears; this fact is in accordance with the explanation given above, the antidromic volley serving to discharge all the motoneurones simultaneously. After the quiet interval belated impulses again succeed in causing excitation.

(7) **Refractory Period.** The neurones, after having discharged an impulse pass through a refractory period, but this is not easily measured with accuracy. The best method is to apply an antidromic stimulus to the motor nerve † (Fig. 164). This causes all the motoneurones to discharge, and it is found that for about 11 msec. afterwards afferent stimuli are ineffective. Hence the refractory phase is assessed at about 11 msec. Longer periods of irresponsiveness in reflexes occur, but are usually attributed to inhibition, *i.e.*, the neurones are not refractory, but are probably not receiving enough afferent impulses.

(8) **Fractionation.** When a given afferent nerve is stimulated with a maximal shock, the reflex contraction of a muscle is only a fraction of what it is capable of producing when its motor nerve is stimulated directly. Thus a given tibialis anticus muscle

---

* Forbes. *Physiol. Rev.*, 1922, **2**, 361.

† Eccles. *Proc. R. S.*, 1931, **107B**, 557.

could exert a tension of 2,160 grammes when directly stimulated. When reflexly stimulated through the internal saphenous nerve the tension was only 800 grammes, when by the musculocutaneous branch of the peroneal, 1,700 grammes; by the external plantar 1,240 grammes; by the internal plantar 1,330 grammes; by the small sciatic 680 grammes, and so on. This phenomenon is called "fractionation." It means that each afferent nerve can only command a portion of the motoneurone pool that supplies the muscle.

It should be noted, however, that the *sum* of the tensions independently obtainable from *all* the afferents greatly exceeds the maximal tension on motor nerve stimulation. In the above example the tensions produced by separately stimulating ten different afferent nerves, added up to 8,370 grammes, whereas the motor tetanus gave 2,160 grammes. This is readily explicable as an example of the *Principle of Convergence*, the areas of influence of the various afferents overlapping considerably, as illustrated in Fig. 165.

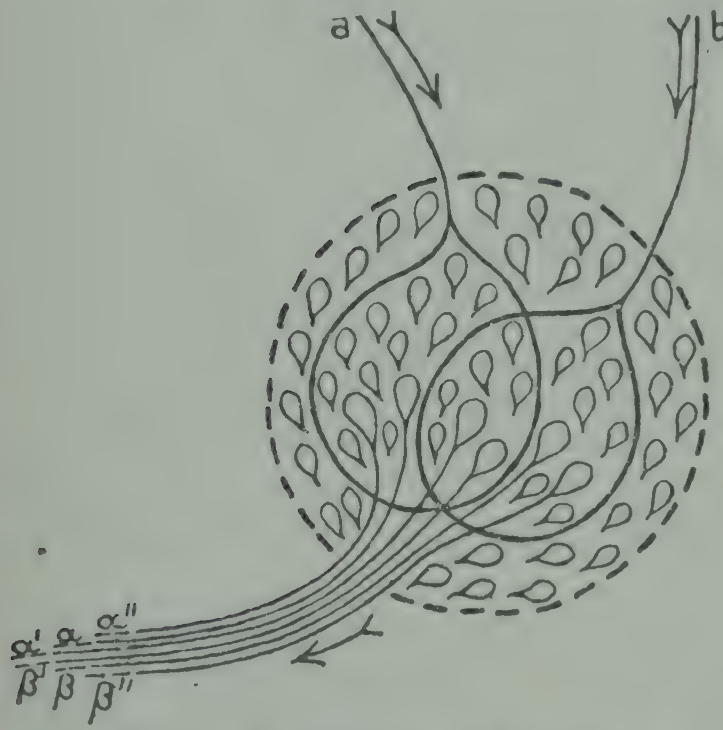


FIG. 165. To illustrate the principle of Convergence and the phenomenon of Occlusion, due to overlap in the neurones activated respectively by the two afferents a and b. The efferent neurones α' and β', for instance, are activated by both afferents. (CREED, DENNY BROWN, ECCLES, LIDDELL and SHERRINGTON. "Reflex Activity of the Spinal Cord," Oxford University Press, 1932.)

(9) Occlusion. The phenomenon called occlusion is another consequence of the principle of convergence. When reflex contraction of a muscle is produced by the simultaneous stimulation of two afferent nerves, the resulting tension is usually definitely less than the sum of the tensions set up by stimulation of the two nerves separately. Thus, when either of two plantar nerves was stimulated, 1,570 grammes tension was produced in a tibialis anticus muscle. If both nerves were stimulated one might therefore expect a tension of 3,140 grammes, but that actually developed was only 1,810 grammes, *i.e.* 1,330 grammes less. This phenomenon is again due to certain of the motor neurones being stimulated by both afferent nerves. Whereas one plantar nerve causes reflex stimulation of x motor neurones, and the other plantar nerve causes the reflex stimulation of y motor neurones the two nerves together do not activate $x + y$ motor neurones because some of those included in x are also included in y, *i.e.* there is considerable overlap, or "occlusion."

Occlusion is greatest in closely allied afferent nerves, because they are more likely to control adjacent parts of the motoneurone pool; it is also greater with more powerful and numerous stimuli, because these extend the area controlled by each afferent and so increase the overlap.

We may now consider some other characteristics which are seen to greater advantage in reflexes other than the flexor.

(10) Inhibition. This can be elegantly demonstrated on the crossed extensor reflex. If the central end of the contralateral sciatic nerve is stimulated by rapid shocks the quadriceps extensor enters into sustained contraction. If, during this tonic state, the central end of an ipsilateral afferent nerve, *e.g.* tibial, be strongly stimulated, the extensor contraction is instantly abolished. In a similar way the flexor reflex which is excited by an ipsilateral afferent can be inhibited by stimulating a contralateral afferent. It will be evident that this phenomenon of central inhibition, *e.g.* of an extension by a flexor reflex, must be closely related to that of reciprocal innervation, and results in a smooth and economical functioning of the muscles; the afferent nerve which causes reflex flexion, at the same time inhibits extension, and *vice versâ*.

The anatomical basis of the various events involved in the carrying out of such a reflex as that just studied is shown in Fig. 166. In this diagram the nerve fibre α represents the nociceptive nerve from the skin of the foot. This passes by a dorsal root into the spinal cord, where it gives off a number of collaterals. Those collaterals, which affect the motoneurones to the two hind limbs, can be divided into two sets, one of which produces excitation in some neurones, whilst the other produces inhibition of the motoneurones to the antagonistic muscles. The single afferent nerve fibre is, therefore, with regard to one set of its central terminal branches, specifically excitor and, in regard to another set, specifically inhibitor. In the case in point the central terminal branches of the nerve α are excitor for the flexor and inhibitor for the extensor muscles of the same side, and the reverse for the muscles of the opposite side.

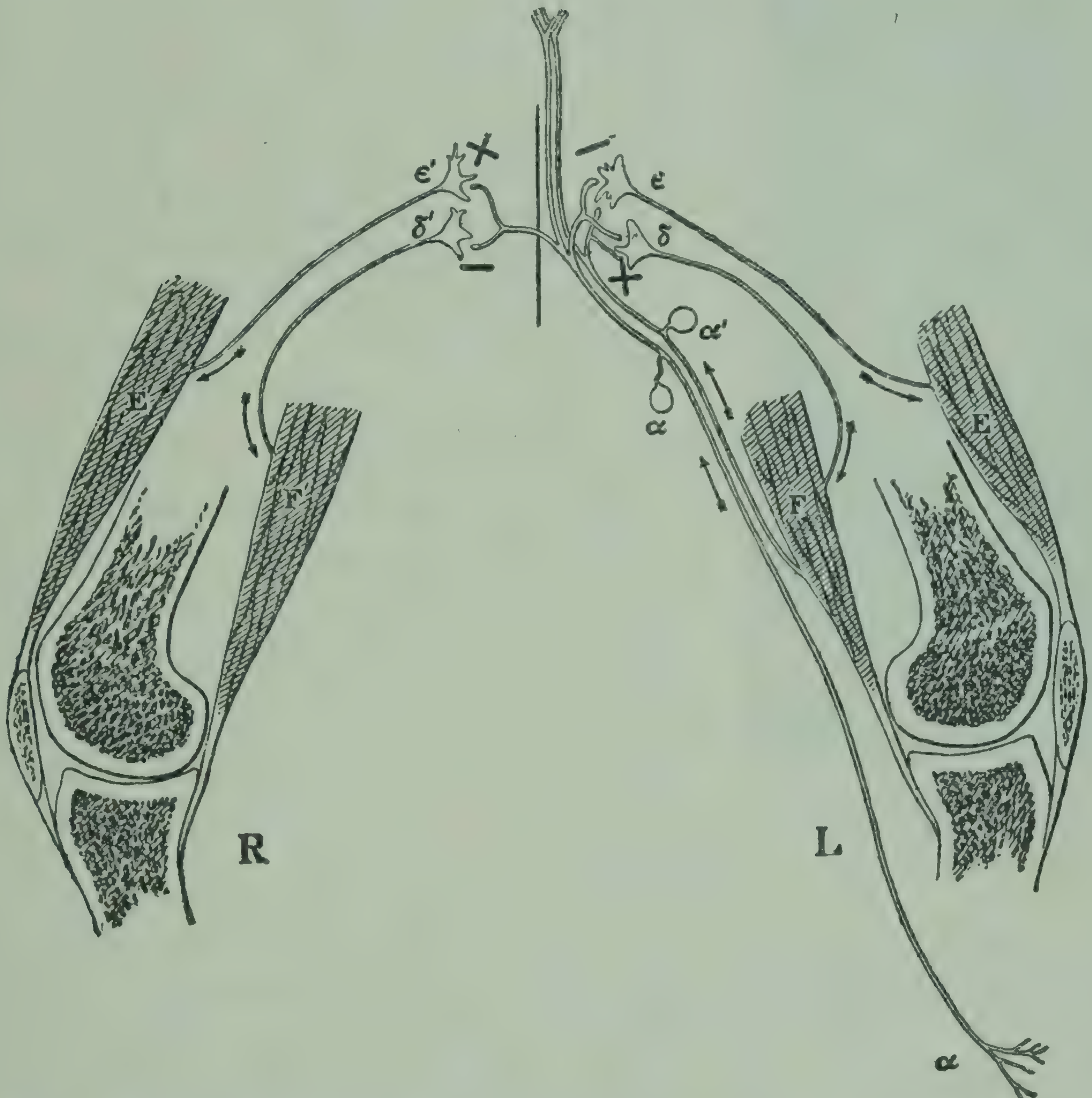


FIG. 166. Diagram indicating Connections and Actions of two Afferent Spinal Root Cells a and a' in regard to their reflex influences on the extensor and flexor muscles of the two knees. The sign + indicates an excitatory effect, the sign — an inhibitory effect. (SHERRINGTON.)

The diagram shows also how the contraction of the flexor muscle itself initiates a secondary reflex process from the muscle up the nerve fibre α' and back again to the muscle by the efferent neurone. This muscular afferent nerve also has central terminations of two signs—excitor to itself and inhibitor to the antagonistic muscles.

It is important to note that the efferent path innervating, say, the flexor muscles is common to many reflexes. It is used, for instance, by mutually exclusive reflexes such as the "scratch reflex" and the "flexor reflex." Evidently the mutual inhibition of opposed reflexes occurs, not in the "final common path"—*i.e.* in the motor neurones, which must always remain open—but farther back in the arc, probably at interneurones near its afferent side.

The biological value of inhibition is clear. For proper co-ordination there must be no clashing between opposing activities of different parts. One stimulus may *reinforce* another if the reactions ensuing on the two stimuli are *allied*—*i.e.* tend to co-operate one with another. But if one part is engaged in any action, this action must be the policy of the body as a whole. Yet the surface of the body is being continually played upon by ever-changing stimuli, tending to excite first one reflex and then another, and the activities so excited would produce confusion in the conduct of the animal, if there were not some means by which at any one time only one reaction should be carried out. The imperative stimulus should dominate the actions of the body as a whole. There may be a struggle of different stimuli, but one must finally be *prepotent* and annul altogether the influence of the others. The study of the spinal reflex shows that this end is obtained by the process of inhibition. Every successful reflex inhibits all other reflexes which are not co-operative with it. We may, for instance, stimulate the area of skin which gives rise to the scratch reflex, and at the same time apply a painful stimulus to the foot. The result is not a movement compounded of the two reflexes, but the scratch reflex immediately comes to an end, and the flexor reflex occurs. When this in its turn stops, the scratch reflex may be once more resumed (Fig. 167).

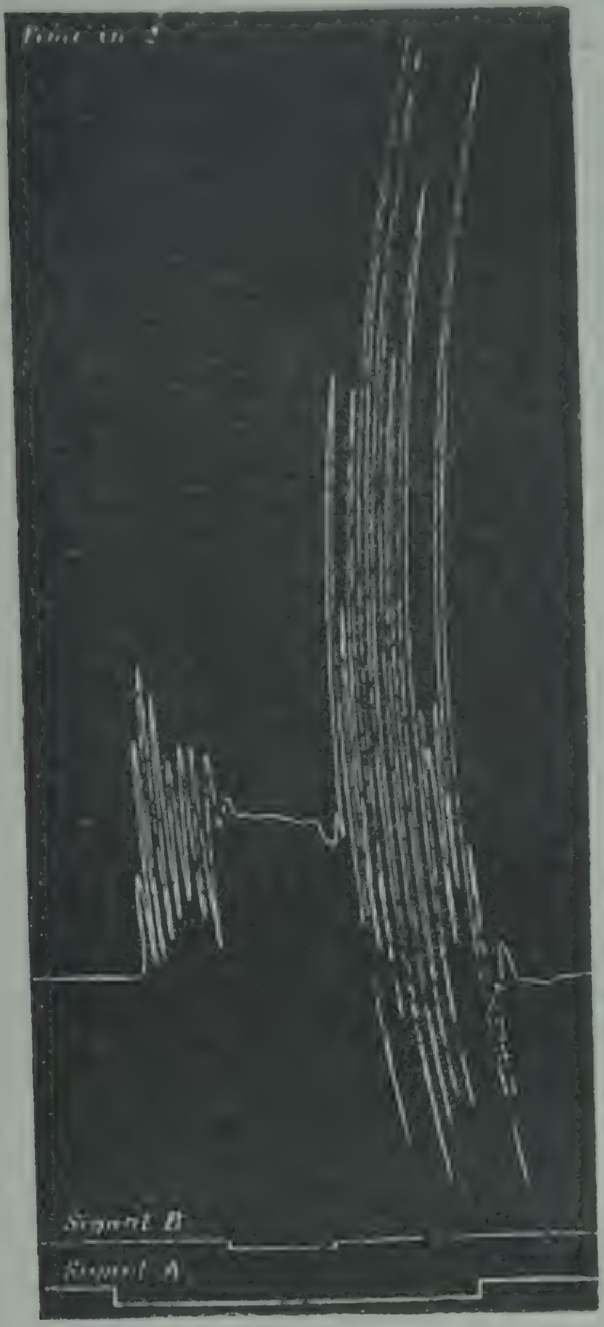


FIG. 167. Scratch Reflex temporarily inhibited by application of a Nociceptive Stimulus to Foot.

Signal A, stimulation of scratch area. Signal B, stimulation of paw by strong induction shock.

Central Inhibitory State. Inhibition is believed to be produced as a result of the establishment in the centres of a condition called the *central inhibitory state* (c.i.s.), which is antagonistic to the c.e.s. As we have already seen, some of the phenomena of inhibition may be due to nothing more than a lack of facilitation. But in other examples, what is called "*direct inhibition*" has been proved to occur. In this, a testing volley (*e.g.* along 6th L. dorsal root) even if of threshold strength * can inhibit a reflex from a stimulus synchronously despatched along a two-neurone arc (*e.g.* from 1st S. dorsal). This means that some positive

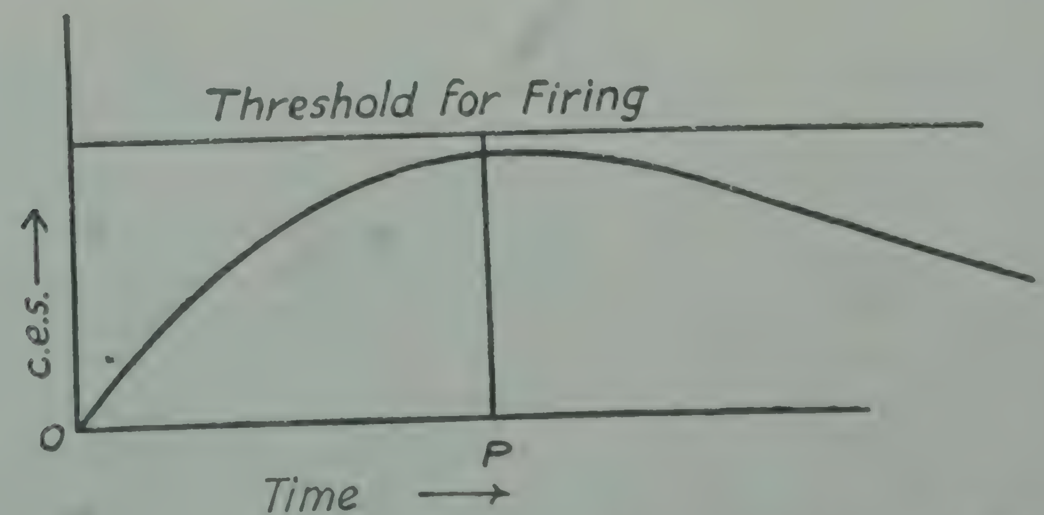


FIG. 168. Probable relation between time and intensity of *c.e.s.* The moment P, when *c.e.s.* is at a maximum, is about 6σ after the arrival of the impulse at O.

---

* LLOYD. *J. Neurophysiol.*, 1941, **4**, 184; 1943, **6**, 111.

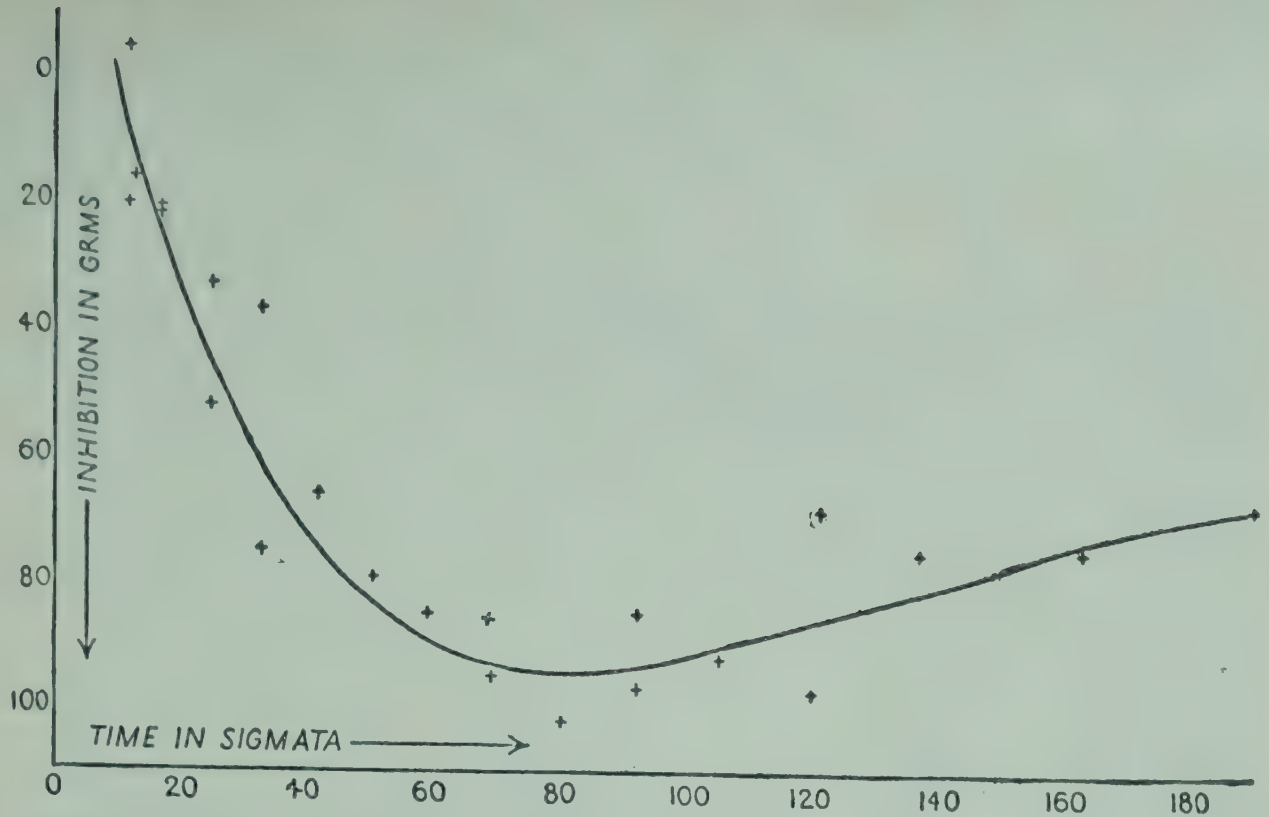


Fig. 169. Tensions of Reflex Twitches (Ordinates) Plotted against Intervals by which an Inhibitory Volley precedes the Excitatory (abscissæ). (Creed *et al.* "Reflex Activity of the Spinal Cord.")

inhibitory effect has been produced on the motor neurones, but the nature of this effect is unknown.

In most instances, inhibition occurs, not at the effector neurones, but at interneurones; its duration is then longer.

If an inhibitory volley *precedes* the excitatory one, the extent of the resulting inhibition depends on the time interval between them. It is maximal (for the inhibition of the flexor reflex) at about 80 msec. (Figs. 168 and 169).

The c.i.s. produced by an afferent stimulus thus rises more slowly than c.e.s. (which is at a maximum in about 6 msec.). It also lasts longer, and then slowly subsides.

The c.i.s. can be summated, either temporarily or spatially. Thus two inhibitory stimuli produce more inhibition than one, if suitably spaced. An antidromic volley cannot, however, discharge c.i.s. as it can c.e.s. Excitation and inhibition can be quantitatively balanced one against the other and the stronger can ultimately overcome the other (Fig. 170).

The exact way in which inhibition is produced has been the subject of many speculations. Many kinds of inhibitory phenomena are met with in the body, and the explanation of one may not be applicable to another. Recent evidence seems to point to the probability that the neurones of the central nervous system are normally exposed to constant subliminal bombardment with impulses and that excitation occurs when new impulses arrive in sufficient numbers and at appropriate instants. Inhibition then comes about if this background of facilitation is withdrawn, because, as we have seen, it needs more than one impulse to excite a motoneurone to the discharge point.

Examples of direct inhibition of motoneurones (*i.e.* like that of vagus on the heart) have been discovered, but cannot at present be explained.

An explanation of reciprocal innervation has been advanced by Gasser on the following lines (Fig. 171). Suppose that two synapses must be excited within 0·5 msec. in order to excite the neurones. When fibre I is repeatedly excited *b* and *c* and hence F, a flexor motoneurone, are stimulated.

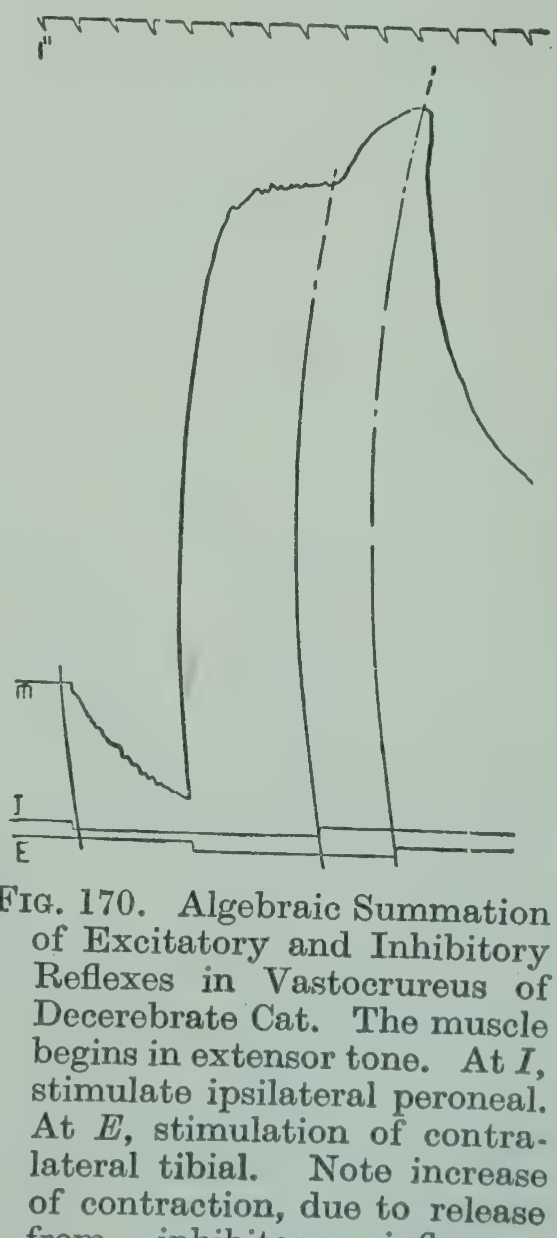


Fig. 170. Algebraic Summation of Excitatory and Inhibitory Reflexes in Vastocrureus of Decerebrate Cat. The muscle begins in extensor tone. At *I*, stimulate ipsilateral peroneal. At *E*, stimulation of contralateral tibial. Note increase of contraction, due to release from inhibitory influence, when *I* ceases. (Creed *et al.* "Reflex Activity of the Spinal Cord.")

Now, suppose that while this is happening, fibre II is stimulated. It has three synaptic knobs on *b* (and two on *a*), and can excite *b* at its own rhythm, which is different from that of I. Neurone *b* now comes under the control of II, and discharges at a different rate from *c*, so that F does not receive successive impulses from *b* and *c*, within 0·5 msec. and therefore ceases to respond. Neurones *a* and *b*, however, act in unison and so excite E, causing contraction of the extensor.

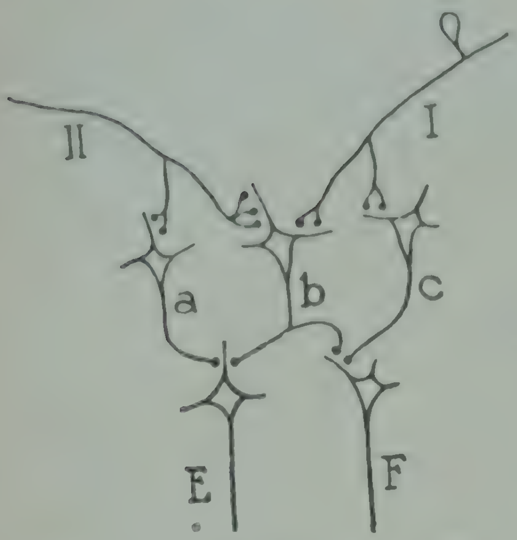


FIG. 171. To illustrate Theory of Central Inhibition and Reciprocal Innervation. (GASSER. Harvey Lecture, 1937.)

One suggestion as to the nature of central inhibition is that it is due to a polarisation change on a different part of the nerve cell. According to this theory,* the arrival of impulses at the cell body or dendrites (the excitatory pole) causes polarisation changes which result in an electrotonic current, as shown in A of Fig. 172. This current, arriving at the axon hillock, may generate an impulse along the axon. But other synapses end on the axon hillock (the inhibitory pole) and these cause a current which opposes the electrotonic current from the excitatory pole, as shown in B, Fig. 172. Excitatory and inhibitory afferent stimuli can thus be balanced one against the other. The theory also accounts for reciprocal inhibition as shown at C, Fig. 172, by supposing that collaterals from the cells of opposing neurons pass each to the inhibitory pole of the other.

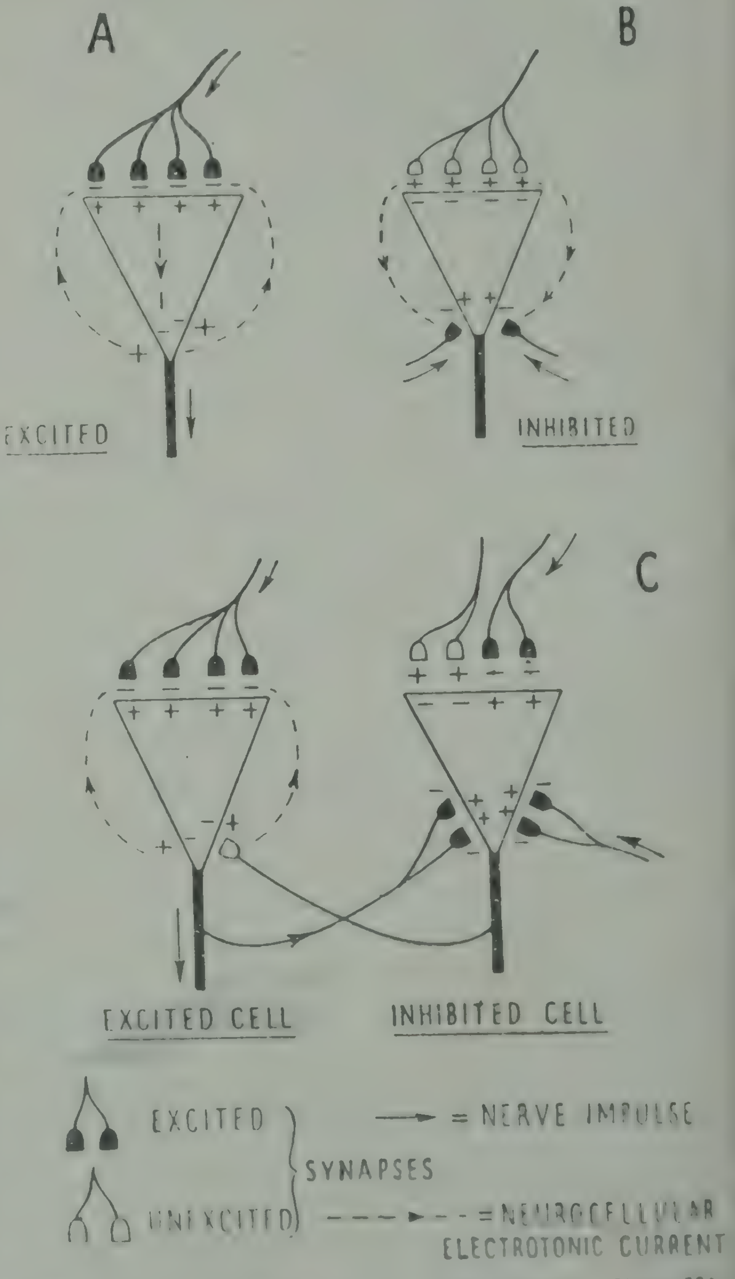


FIG. 172. Diagram to illustrate Gesell's theory of central inhibition.

(11) "Rebound." Inhibition is often followed by a condition of increased excitability, as a sort of swing-back of the state of excitation. Thus, suppose the scratch-reflex is excited and then, while the excitation is still continued, the reaction is inhibited by excitation of the flexor reflex. As soon as the flexor reflex has passed off, the scratch reflex returns with an intensity greater than before (Fig. 167, p. 258). It may also be seen sometimes as an increased after-discharge following the cessation of an excitatory stimulus. In the knee-jerk a succession of inhibitions and rebounds may result in a clonic contraction.

(12) Negative Successive Induction. When a reflex subsides, it often tends to be followed by an opposing movement, a phenomenon called *negative successive induction*. This explains the tendency of the spinal cord to an alternation of response; every act tending as a result of negative spinal induction to induce the antagonistic act. Thus if a spinal dog be held up vertically, so that the hind limbs hang freely, these latter execute a series of alternate movements of flexion and extension—the "*mark-time*" *reflex*. The starting point of these is the stretching of the anterior thigh muscles. Once started they continue of themselves, each act exciting the alternating antagonistic act (Fig. 173).

---

* GESELL. *Ergebn d. Physiol.*, 1940, **43**, 477; *Amer. J. Physiol.*, 1947, **148** 515.

The movements of biting and mastication are probably largely effected by a series of rebounds aided by negative successive induction, originated by the mechanical stimulus of food in the mouth.

(13) Recruitment. Many reflexes, especially the extensor ones, show a much slower development than does the flexor response or that obtained by stimulation of the motor nerve; this is due to the fact that the motor units are not thrown into action all at once, but are brought into play in regularly increasing numbers up to the maximum. The decline of these reflexes is also gradual. This phenomenon is called *recruitment*, and is responsible for the steady rise and fall of contraction in certain natural movements, such as those of the respiratory muscles.

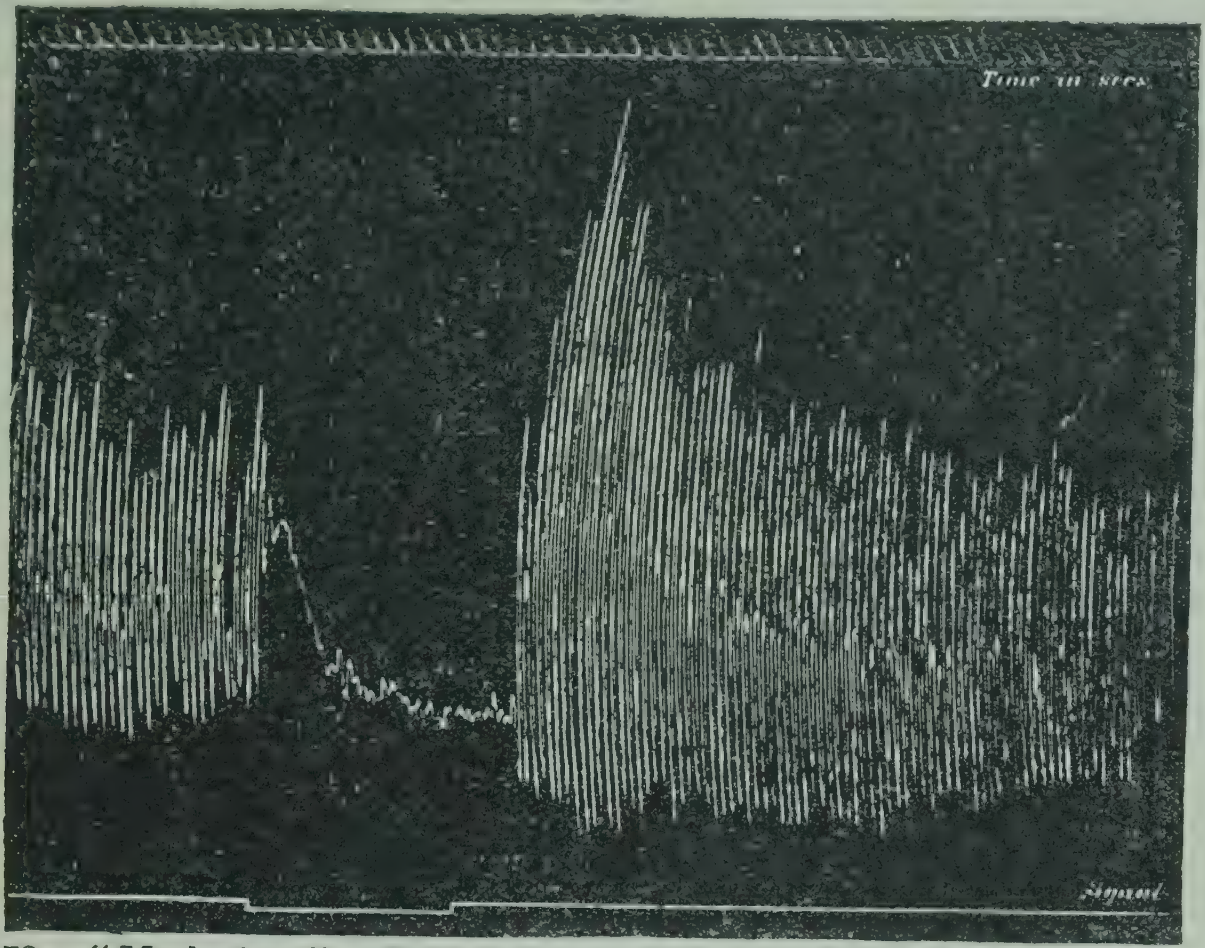


FIG. 173. "Mark-time" reflex in spinal dog, inhibited by slight stimulation of the tail (duration of stimulation shown by signal). Note the augmentation of the mark-time reflex following the inhibition (*rebound*). (SHERRINGTON.)

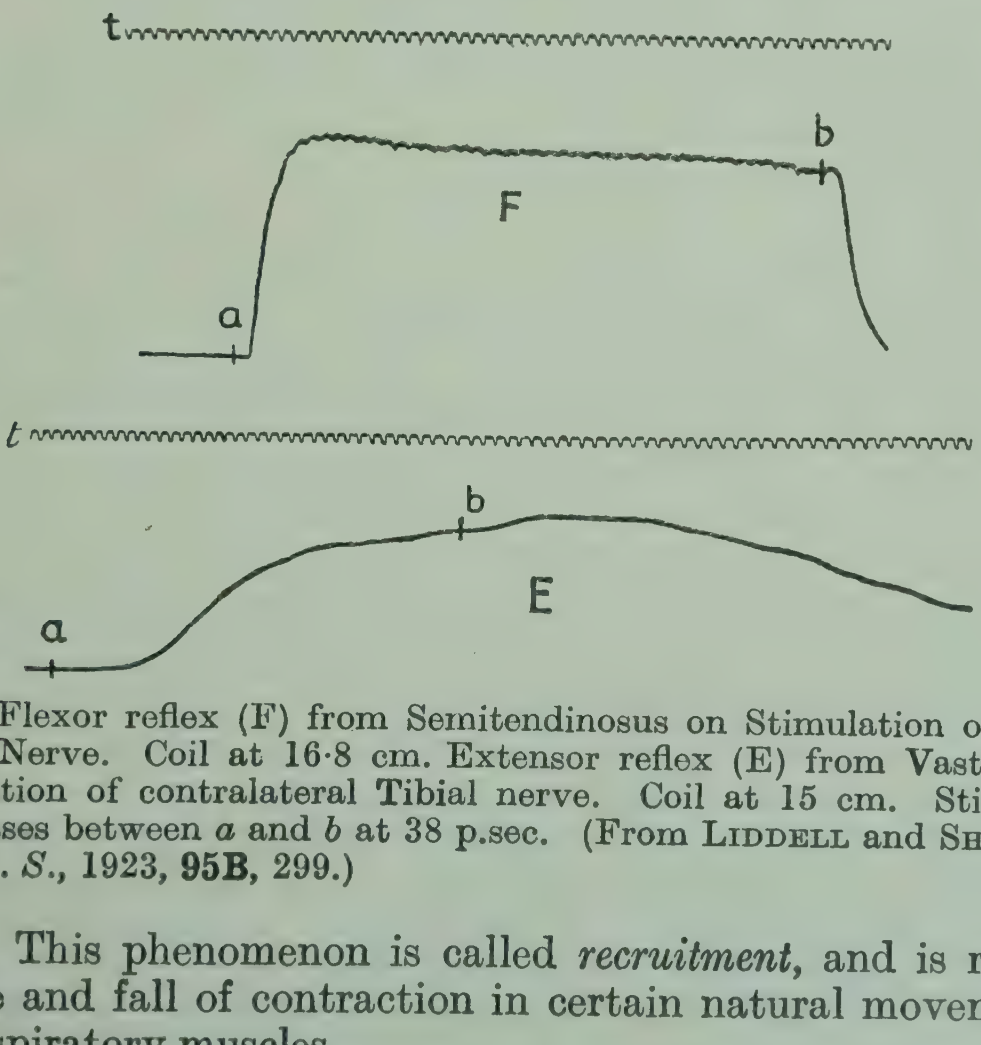


FIG. 174. Flexor reflex (F) from Semitendinosus on Stimulation of Ipsilateral Tibial Nerve. Coil at 16·8 cm. Extensor reflex (E) from Vastocrureus on stimulation of contralateral Tibial nerve. Coil at 15 cm. Stimulation in both cases between *a* and *b* at 38 p.sec. (From LIDDELL and SHERRINGTON. *Proc. R. S.*, 1923, **95B**, 299.)

When, as in many flexor reflexes, all the motor units employed are thrown into use at the commencement of the response, the onset is sudden, and is of the type called by Sherrington "*d'emblée*" reflexes. Fig. 174 shows the striking difference between flexor and extensor reflexes.

It will be seen that the flexor response shows less delay, needs a weaker stimulus, has a more abrupt rise and fall, has little after-discharge, and shows the rhythm of the afferent stimuli. The slow falling-off of the extensor response is clearly due to prolonged after-discharge. Its slow rise, or recruitment, is probably explicable as due to a larger number of internuncial neurones in the reflex arc. Hence the delay in commencement is longer than for the more direct flexor reflex, because the c.e.s. takes longer to work up to an excitatory level; it also more slowly reaches a maximum as the later impulses trickle into the motoneurone pool. Also, owing to the many different delay paths the rhythm of the afferent stimulus is damped out, so that a smooth contraction results even when the rate of afferent stimulation is relatively slow.

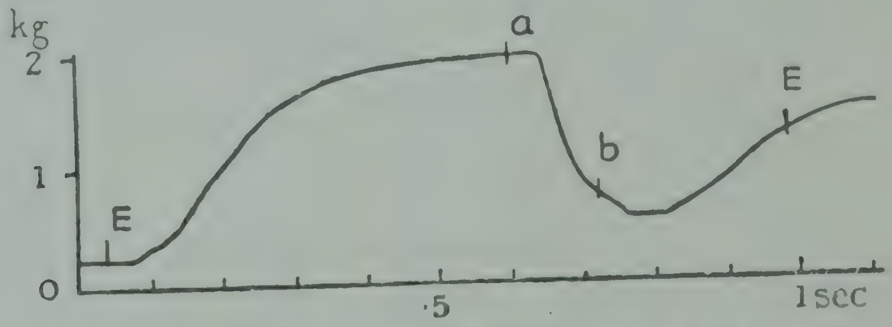


FIG. 175. Recruitment and "after-discharge" of Inhibition. *Rectus femoris.* E to E, stimulation of contralateral afferent causing crossed extensor reflex. *a* to *b*, stimulation of ipsilateral afferent causing inhibition of slow onset and outlasting the stimulation. (LIDDELL and SHERRINGTON, *Proc. Roy. Soc.*, 1925, **97B**, 488.)

Inhibition, as well as excitation, shows the phenomenon of recruitment, if the inhibitory stimulus is weak, *i.e.* the motoneurones cease working in succession, not all at once (Fig. 175); it also shows after-discharge.

There is an alternative explanation of recruitment. The extensors in the decerebrate preparation are normally in a state of tone, due to stretch-reflexes. When an extensor contraction begins, the stretch is relieved, so that the rise of tension is slowed down: also, afferent impulses having an inhibitory effect pass up from the muscle, and these have the same effect of acting as a brake on the rise of tension. The converse happens when relaxation occurs. The evidence for this view is that if the afferents from the extensor muscles are divided, the characteristic recruitment and after-discharge are much lessened, and hence they must be partly proprioceptive in origin.*

The extensor muscles therefore tend to give smooth contractions and steady relaxations.

(14) **Fatigue.** If a given reflex be repeatedly elicited, the reaction becomes feebler and finally disappears altogether (*e.g.* the flexor reflex in 30 seconds) long before any signs of fatigue in the motor apparatus can be detected on stimulation of the motor nerve. Since we know that it is practically impossible to fatigue nerve fibres, we must conclude that the seat of fatigue is in the grey matter of the spinal cord itself.

Central fatigue might be due to a variety of causes, *e.g.* depletion of energy-yielding materials in the soma, depletion of sources of humoral transmitters at synapse, accumulation of substances causing block at the synapses, or failure of normal electrical polarisation at the synaptic membrane. The first of these can be discounted because the fatigue changes are strictly localised at the synaptic endings. This is shown by the fact that only one path of approach may be fatigued: thus after the flexor reflex has been fatigued by prolonged stimulation of the popliteal nerve, it can still be elicited by stimulation of the peroneal nerve, and *vice versâ*.†

---

* SHERRINGTON. *Q. J. Exp. Physiol.*, 1909, **2**, 109.

† FORBES. *Am. J. Physiol.*, 1912, **31**, 102.

The phasic reflexes, such as the flexor reflex, are much more readily fatiguable than are the tonic reflexes, such as the extensor and postural ones. These latter also are evoked from less quickly adapting sensory receptors, and are better equipped for sustained activity. We may suppose that this is possible without synaptic fatigue because the synapses function in rotation, and have time to recover between their spells of action.

(15) " Block " or Resistance. As we have seen, there is an absolute block to the passage of an impulse backwards through a synapse. The phenomena of fatigue show that there may be resistance even to the passage of an impulse in the normal direction. When we study the structure of the central nervous system, we find that although there are certain shortest possible paths for every impulse, yet so extensive is the branching of the entering nerve fibres and so complex are the neurone systems with which they come into connection, that if all resistance were removed, an impulse entering along any given fibre could spread to practically every neurone in the spinal cord and brain. Such a result is indeed observed in animals poisoned by strychnine. In these, the slightest stimulus applied to any part of the skin excites strong tonic spasms in the musculature of the whole body. Every sensory nerve fibre, that is to say, can discharge into every motor neurone of the cord. That this result does not ensue on localised stimulation in a normal animal is because of the varying resistance to the passage of an impulse into the several neurones with which the entrant fibre comes into anatomical relation. A small stimulus will discharge only along the few neurones where the resistance is lowest. Increase of the stimulus, either by adding to its strength or by summation of weak stimuli, will enable the impulse to irradiate and therefore will elicit a more widespread response. Only when the " blocks " are entirely removed by the administration of strychnine, or when the stimuli are abnormally powerful and long continued will the impulse spread to all regions of the central nervous system, so that the response becomes general and inco-ordinate.

STRETCH REFLEXES

Normal muscles are in a state of slight constant tension, or *tone*, and in the decerebrate animal this is greatly exaggerated in the extensor muscles. Section of either the motor or the afferent nerves of the muscle destroys this tone and the muscle becomes limp. When the tendon of such a denervated muscle is pulled upon, the muscle is merely passively stretched. It is quite different if a pull is given to the tendon of an innervated extensor muscle, such as the quadriceps, of a decerebrate preparation. The muscle in such a case returns pull for pull, actively contracting against and in proportion to the stretching forces. This response is a *stretch reflex*. Such an experiment is illustrated by Fig. 176. The quadriceps of a decerebrate cat was attached by the patellar tendon to an isometric lever and was subjected to stretching by pulling the animal in the opposite direction to the lever by lowering the animal table by a few millimetres.

A tension of nearly 4 kg. resulted, but after denervation the passive lengthening of the muscle by the same stretch gave only a little over 1 kg. tension. The reflex and proprioceptive character of the pull is demonstrated by the fact that it disappears completely if the afferent fibres from the muscle are divided by section of the appropriate dorsal nerve roots.

The stretch reflexes are shown especially well in the extensors of decerebrate preparations ; sustained stretch reactions are not seen in flexors. These proprioceptive reflexes originate in the stimulation of the stretch receptors in the muscle and tendon. As we have already seen (Chapter XIV) the stretch receptors start off trains of impulses when the muscle is stretched, and as the stretch increases the impulses follow each other more frequently, and along more and more fibres. Muscle stretch receptors do not adapt rapidly, and so keep up trains of impulses along the afferent fibres from the muscle as long as the stretch is maintained, and if this is increased, more and more motoneurones are recruited. The contraction is therefore as sustained and as strong as the stretching force, but only affects the muscle, or that part of it, subjected to stretch—it never spreads, as flexor reflexes do. Indeed, it is probable that each muscle spindle serves as the starting point for stretch reflexes confined to immediately neighbouring fibres.

The pain receptors, by way of contrast, adapt quickly, and so are more suited to evoke the brisk but more evanescent flexor responses.

Stretch reflexes are mediated by large afferent fibres, are effected by simple two-neurone arcs, and therefore have short latent periods (6 to 20 msec.) and show little after-discharge. They can at once be inhibited, as Fig. 176 shows, by exciting an opponent reflex. The rate of discharge of the neurones is slow, *e.g.* about 8 p. sec., asynchronous, and involves little metabolic cost. It is probably on this account that fatigue does not occur.

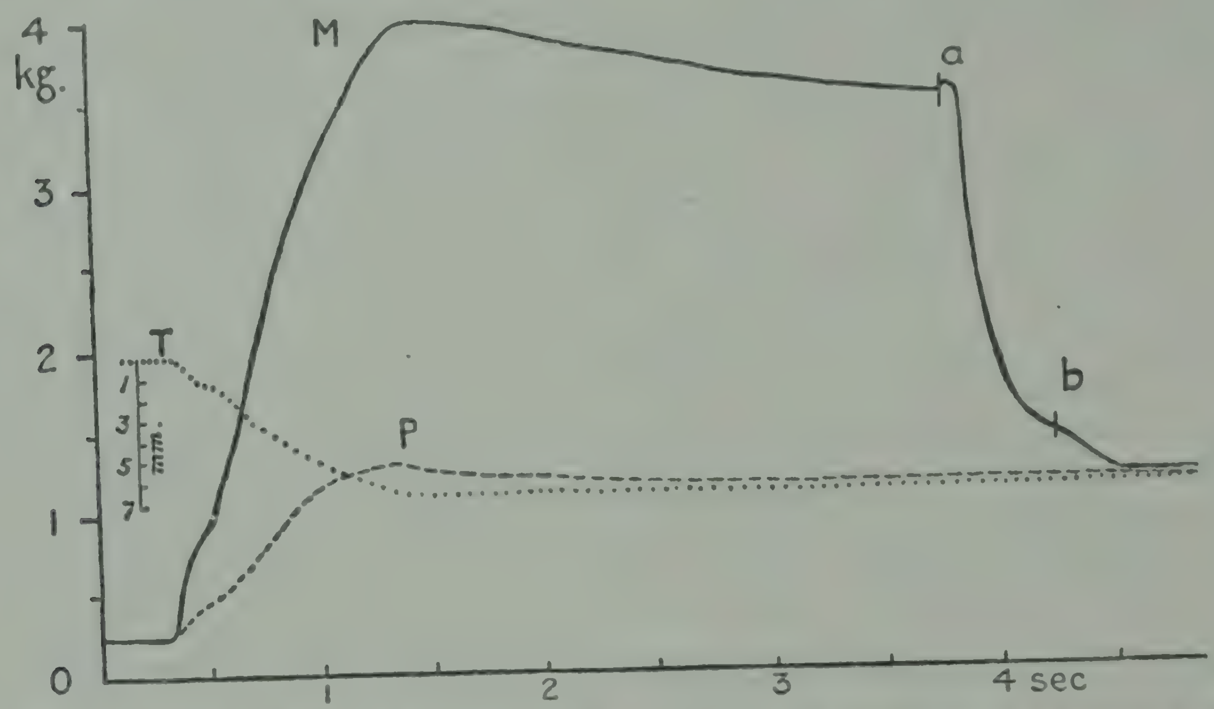


FIG. 176. Stretch Reflex of *Quadriceps*, decerebrate Cat. Extent of stretch shown by curve T; M, tension of muscle; P, passive tension by repeating stretch after denervation of muscle. *a*, *b*, inhibition by stimulation of ipsilateral tibial nerve. (LIDDELL and SHERRINGTON, *Proc. Roy. Soc.*, 1924, **96B**, 212.)

The stretch-reflexes have great importance in connection with the maintenance of posture, of which they form the background. They are seen in those muscles that normally counteract the effect of gravity, and their tonic nature enables an animal to maintain its normal posture without fatigue. When an animal stands erect, gravity tends to cause flexion at hip, knee and ankle, but any tendency of the limbs to give way is instantly corrected by the stretch-reflexes, which convert the limbs into rigid pillars.

Types of Muscle. Most of the pure flexor muscles are made up of pale fibres, the rapid contraction of which is suited to their function. Extensor muscles, on the other hand, are capable both of slow, sustained contractions, as in stretch reflexes, and of quick ones, and accordingly consist of two parts, a quick acting part, of pale muscle, and a slow one of red muscle.* Thus, soleus and gastrocnemius extend the ankle, the former slowly, the latter quickly; medial short head of triceps extends the elbow slowly and the lateral short head quickly, and so on. Physiologically they are distinct muscles, though often associated anatomically as separate "heads." Usually the slower muscle is the more deeply placed.

Tendon-reflexes. The knee-jerk, and other examples of "tendon-reflexes" already referred to in Chapter XIII, are special cases of the stretch reflex. In them the tug on the muscle is very brief and the response is brisk. They originate from the stretch of the muscle and not from the tendon, since they can still be elicited after cocainising the tendon. The latent period is very short (6 to 8 msec.) because a synchronous volley is sent up from the muscle, and because the arc is short.

---

* DENNY-BROWN. *Proc. Roy. Soc.*, 1929, **104B**, 371.

The electric potential changes during a tendon-jerk show a burst of impulses immediately preceding the contraction, followed by a sudden cessation—the "silent interval" at its height, and then a further train during and after the peak. The silent period has been shown to be partly due to the relief of tension on the stretch receptors when contraction has occurred, partly to reflex inhibition, but mainly to the fact that since the motoneurones are put into action synchronously they naturally all reach the state of lowered excitability at about the same time. The tendon-jerk can be abolished by inhibitory excitation, *e.g.* the knee-jerk by stimulation of an ipsilateral afferent.

The reason why tendon-jerks and stretch reflexes are more easily shown in the decerebrate preparation is probably that there is a constant state of extensor excitation facilitating the responses. In man the knee-jerk is facilitated, or "reinforced" by clenching the fists.

Lengthening and Shortening Reactions. If one attempts forcibly to flex one of the rigidly extended limbs of a decerebrate preparation, the resistance due to the stretch reflex is at first encountered, but if the pressure is continued, the limb quite suddenly folds up—"*clasp-knife reaction.*" This sudden yielding in the face of threatened rupture is a protective response, and is called the *lengthening reaction.* It is due to the fact that at a certain degree of tension inhibitor afferents gain the ascendency. Accompanying this inhibition is a tonic increase of extension in the opposite limb, which is called Philippson's reflex (Fig. 177). It is akin to the crossed extensor reflex.

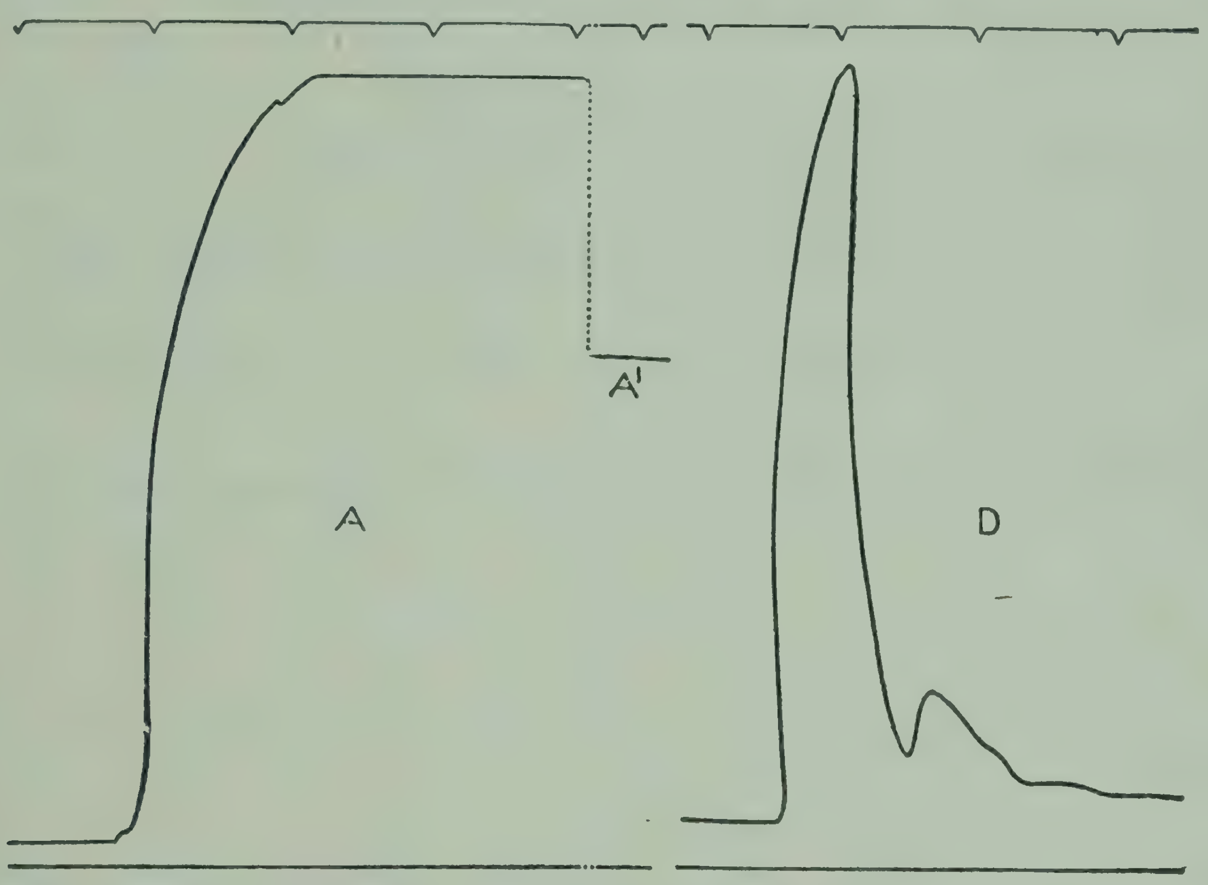


FIG. 177. A, Philippson's Reflex, a Contraction of the Extensors of a Decerebrate Preparation when the opposite Limb is forcibly bent at the Knee. A¹, the same 1 minute later. D, the response 48 days after section of afferent roots of the muscle. Time = secs.

Like the crossed extensor reflex it is converted from a tonic to a phasic response if the extended muscle is first deafferented (Fig. 177, D). This shows that in some manner the extensor reflex is maintained in a tonic state as a result of afferent impulses from the extended muscle—possibly from its own stretch receptors.

The converse of the lengthening reaction is the *shortening reaction* seen if, after passive flexion, the limb is passively re-extended, when it "stays put" in extensor tonus as at the start. By these two reactions a certain plasticity is imparted to the exaggerated postural tonus of decerebrate rigidity.

There is evidence that when a muscle is stretched two opposing reflexes compete. The tendon receptors originate impulses which have an inhibitory effect on contralateral extension, while the annulo-spiral spindle endings of muscle have an excitatory one, provoking the Philippson reflex.*

IPSILATERAL EXTENSOR REFLEXES

Although the commonest type of reflex elicited by the artificial stimulation of afferent nerves consists in ipsilateral flexion, with contralateral extension, there are, nevertheless, various reflexes of ipsilateral extension. Some of these play an important part in the natural co-ordination of reflexes due to the normal stimuli to receptors. The stretch reflex, and the tendon-reflexes are examples of this: the extensor thrust and the positive supporting reactions are others.

---

* MORSON and PHILLIPS. *J. Physiol.*, 1937, **88**, 199.

An ipsilateral extensor reflex is also sometimes obtained on artificial stimulation of an afferent nerve, particularly of one placed distally on the limb. In these circumstances, the contraction of the extensor muscle precedes the normal inhibitory response accompanying the flexor reflex. It is believed that this extensor contraction competes, usually unsuccessfully, with the more powerful inhibitory one, so that it is only sometimes evident, and then often only for a time.

It would seem that each afferent nerve can excite two opposing reflexes, of which one is normally the stronger; the other is called a "*concealed reflex.*" The motoneurones are always subjected to continual processes of excitation and inhibition, and their activity at any moment represents the algebraical sum of these opposed influences.

Reversal of Effect. Owing to the existence of concealed reflexes, it may happen that stimulation of an afferent gives first one reflex and then the reverse one. This may happen especially if the strength of stimulation is altered, or if the stimulation is prolonged, or the conditions changed, as by drugs. Terminal rebound is often such an instance, and possibly the rhythmic reflexes, such as stepping, the scratch reflex, etc., are others.

It was inferred by Graham Brown * that the spinal reflex centre for a limb consists of two "half-centres," one flexor and one extensor, and each afferent fibre reaching the centre divides into branches, some of which go to each half-centre.

For Reference

CREED, DENNY BROWN, ECCLES, LIDDELL and SHERRINGTON (1932). Reflex Activity of the Spinal Cord. (Oxford Press.)

FORBES (1922). The Interpretation of Spinal Reflexes, etc. *Physiol. Rev.*, **2**, 361.

FULTON (1926). Muscular Contraction and the Reflex Control of Movement. (London: Baillière, Tindall and Cox.)

LLOYD (1944). Functional Organization of the Spinal Cord. *Physiol. Rev.*, **24**, 1.

NEWTON (1939). Ch. XI. The Spinal Reflex. Recent Advances in Physiology. (London: Churchill.)

RANSON and HINSEY (1931). *Arch. Neurol. Psych.*, **26**, 247.

SHERRINGTON (1906). The Integrative Action of the Nervous System. (London: Constable.)

* GRAHAM BROWN. *Q. J. Exp. Physiol.*, 1911, **4**, 331; 1912, **5**, 237.

CHAPTER XVIII

THE LABYRINTH, AND POSTURAL REFLEXES

Although quite complex reflexes can be carried out through the mediation of the spinal cord, most of these reflexes are, in the intact animal, considerably modified owing to the action of the higher parts of the central nervous system. This process of modification consists of augmentation of some, and suppression of other reflexes, and a general grouping together of reflexes so that a co-ordinated response of greater complexity results. As an example of this kind of higher function we shall now consider the reflex aspects of posture maintenance. Although this involves the final employment of spinal arcs, the proper association of the various factors which are concerned is carried out largely by centres in the brain stem.

The various movements of the normal body are effected against a background of muscular tonus, which maintains the various postures. The postural tonus of the musculature is effected reflexly, being built up from afferent impulses from two chief sources, viz. the *labyrinths*, and the *muscle proprioceptors*, and from various subsidiary sources. This state of tone is suitably modified when any movement takes place. Many of the movements of normal animals are effected reflexly, and their study provides a good example of the manner in which elementary reflexes, carried out by lower centres, are modified and integrated by the intervention of higher centres of activity.

It is interesting to note that the pathways and arcs responsible for postural maintenance are rapidly conducting two-neurone arcs, whereas those concerned with flexor-reflexes and voluntary action are slow-conducting multi-neurone arcs. Probably this ensures that the initial postural state provides a suitable foundation for the movement.

THE LABYRINTH

The primitive auditory sac arises as a simple involution of the surface. In the course of development the anterior part is modified to form the canal of the cochlea, which is set apart entirely for hearing. From the posterior part there are formed two sacs—the saccule and utricle—and the three semicircular canals. The saccule and the utricle, which receive each a large branch

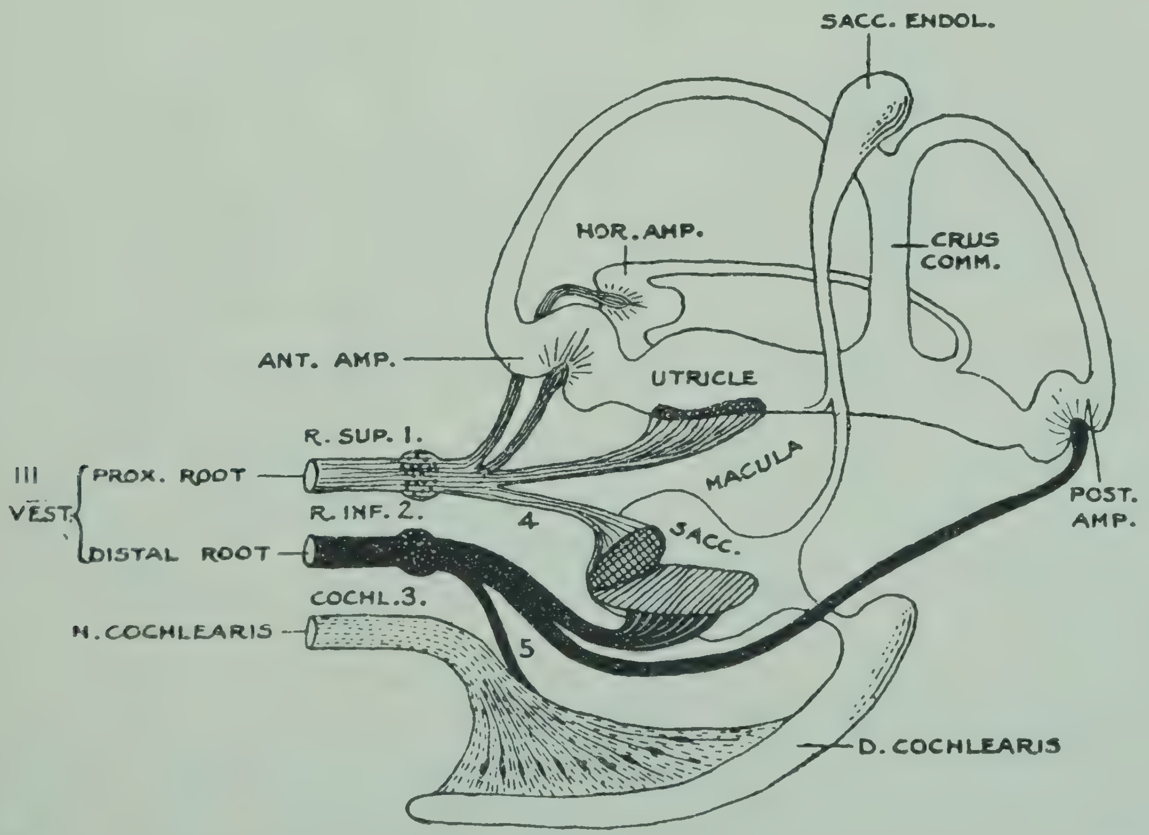


Fig. 178. Scheme of the Innervation of the Membranous Labyrinth. (H. M. de Burlet, *Anat. Anzeiger*, 1924.)

of the vestibular division of the VIIIth nerve, represent the *otolith organ*, which is found in almost all classes of animals.

In mammals, that part of the labyrinth which represents the primitive otolith organs consists of a bony casing containing perilymph, in which is contained the membranous labyrinth, filled with endolymph. The osseous labyrinth consists of a cavity, the vestibule, into which open behind the three bony semicircular canals. In the vestibule are contained two little membranous sacs, the utricle and saccule, the cavities of which are connected by means of the *ductus endolymphaticus* (Fig. 178).

Into the utricle, by five openings, open the three semicircular canals. These semicircular canals are arranged in three planes, each of which is at right angles to the other two, so that the three planes of space are represented. We may distinguish on each side an external or horizontal canal, an anterior vertical canal, and a posterior vertical canal. The two external canals on the two sides lie always exactly in the same plane, which in the normal position of the head slopes backwards and downwards at 30° to the horizontal. Each posterior vertical canal lies in a plane which is parallel to that of the anterior vertical canal of the opposite side. We thus see that these semicircular canals form together three planes, one nearly horizontal and two vertical, the two latter being at right angles to one another (Fig. 179). The membranous canal lies within the osseous canal, a considerable space intervening between the two canals. At one end each of the osseous canals is dilated and the membranous canal undergoes a corresponding dilatation so as to fill up the whole bony canal. In this dilatation, which is known as the *ampulla*, we find the ending of a branch of the vestibular nerve in a special sensory epithelium forming the *crista acustica* (Fig. 180). The crista is composed of hair cells with sustentacular cells between them. The fibres of the vestibular nerve end in arborisations among the hair cells, the hairs of which project into a dome-shaped mucilaginous body, the *cupola*, which lies in the endolymph filling the ampulla (Fig. 180).

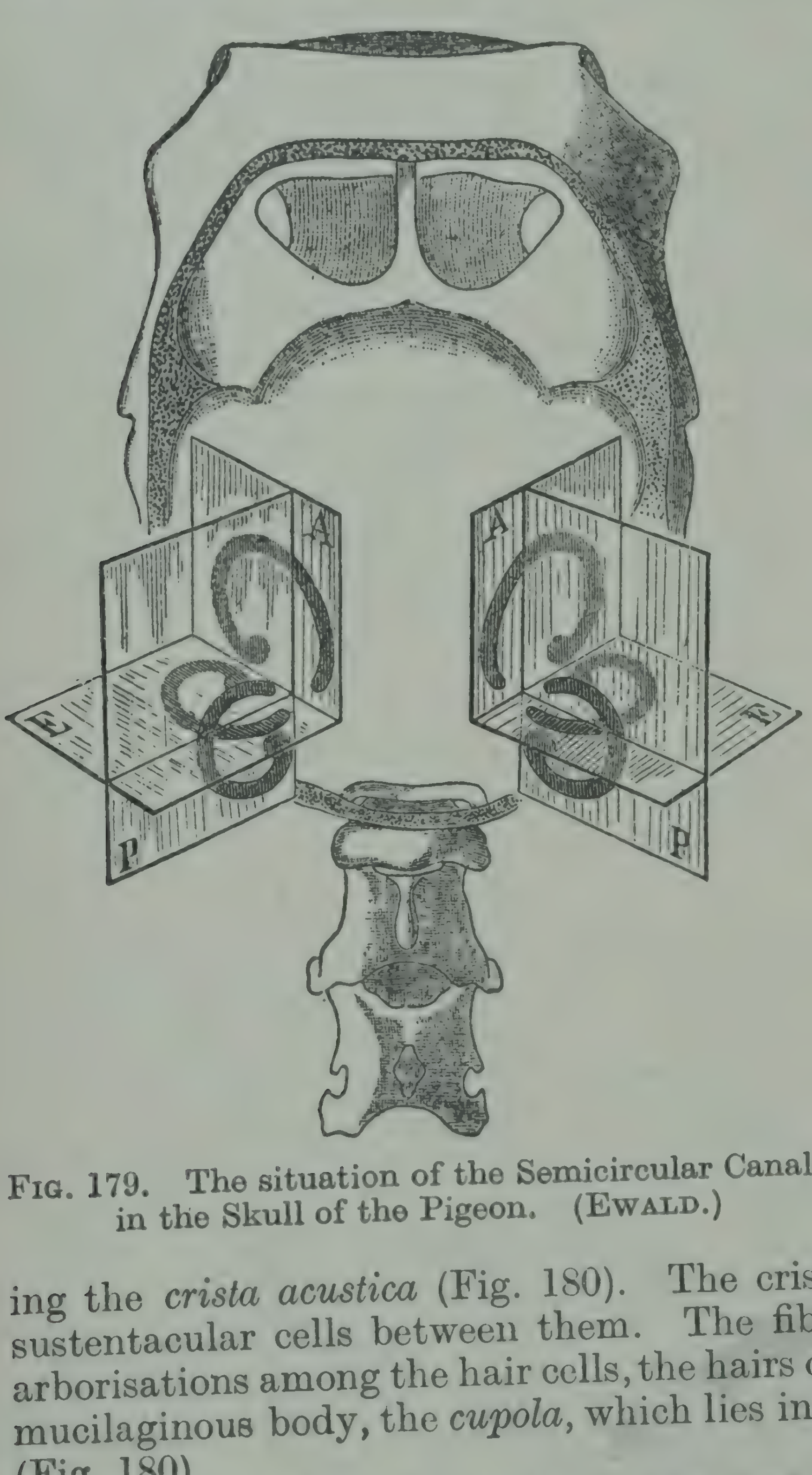
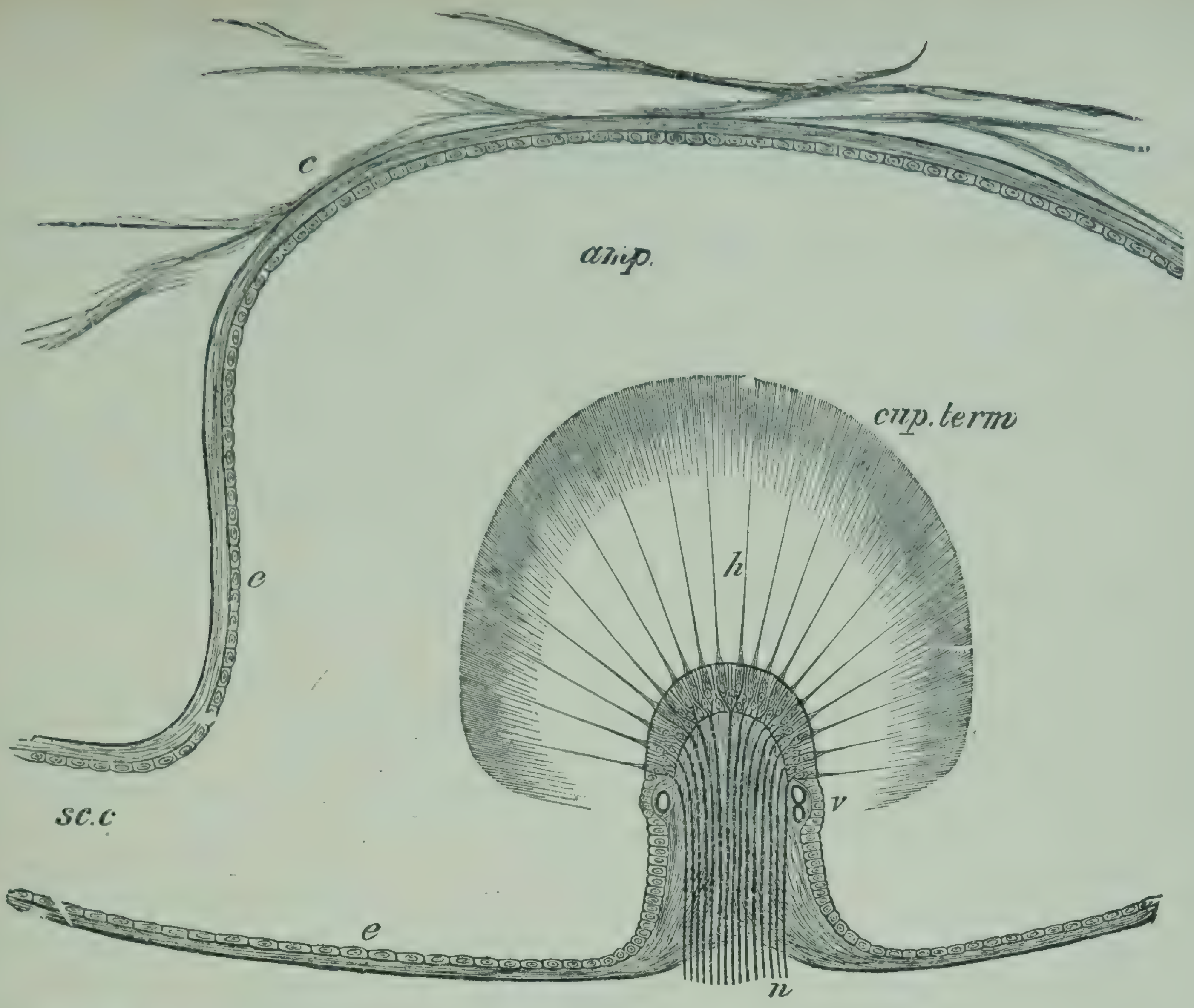


FIG. 179. The situation of the Semicircular Canals in the Skull of the Pigeon. (EWALD.)

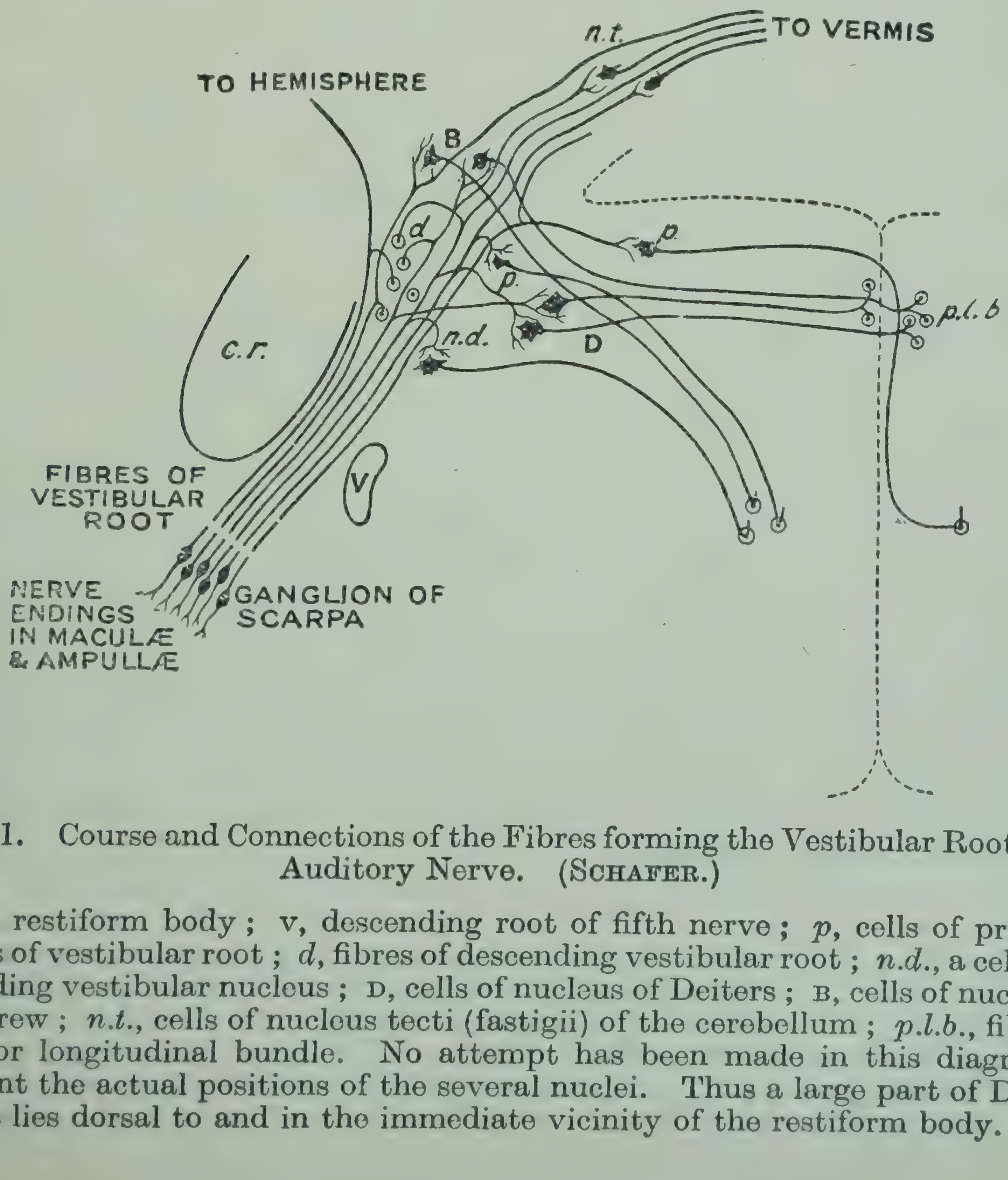
In the utricle and saccule we also find in each an oval patch of modified epithelium. These are special receptor organs, known as the *maculæ acusticæ*, the structure of which is very similar to that of the crista in the ampullæ. In the mucilage covering the hairs of the macula, however, are found numerous small calcareous concretions, the *otoliths*. The macula of the utricle is in a horizontal plane, that of the saccule vertical.

NERVE PATHS CONCERNED. The vestibular part of the VIIIth cranial nerve has its cells of origin in Scarpa's ganglion. It enters the pons abov the cochlear division and passes between the inferior cerebellar peduncl and the descending spinal root of the Vth cranial nerve. Its fibres end b



c
amp.
cup.term
h
e
v
sc.c
e
n

Fig. 180. End Organ (*Crista Acustica*) of Vestibular Nerve in Ampulla of Semicircular Canal, Cup. term., the *cupola*.



n.t.
TO VERMIS
TO HEMISPHERE
B
d
p
p
p.l.b
c.r.
n.d.
D
V
FIBRES OF VESTIBULAR ROOT
NERVE ENDINGS IN MACULÆ & AMPULLÆ
GANGLION OF SCARPA

Fig. 181. Course and Connections of the Fibres forming the Vestibular Root of the Auditory Nerve. (Schafer.)

c.r., restiform body ; v, descending root of fifth nerve ; *p*, cells of principal nucleus of vestibular root ; *d*, fibres of descending vestibular root ; *n.d.*, a cell of the descending vestibular nucleus ; D, cells of nucleus of Deiters ; B, cells of nucleus of Bechterew ; *n.t.*, cells of nucleus tecti (fastigii) of the cerebellum ; *p.l.b.*, fibres of posterior longitudinal bundle. No attempt has been made in this diagram to represent the actual positions of the several nuclei. Thus a large part of Deiters' nucleus lies dorsal to and in the immediate vicinity of the restiform body.

arborising round one of four different collections of nerve cells of which the lateral is called "Deiters' Nucleus." Fibres originating in these nerve cells form four pathways which convey impulses of importance in movements of the eyes, neck and trunk; in addition, large numbers of them go to the cerebellum (Fig. 181).

These paths are (i.) the *posterior longitudinal bundle*, to the oculomotor nuclei. These fibres, concerned with the co-ordinated movements of the eye muscles, run principally from the superior vestibular nucleus ("Bechterew's N.") and from Deiters' Nucleus; (ii.) the *vestibulo-spinal tract*, descends to spinal motoneurones; (iii.) the *vestibulo-cerebellar tract*.

Fibres conveying impulses to the cerebellum run principally from Deiters' Nucleus, but some go direct; all of them enter the inferior cerebellar peduncle, relay in the n. fastigii, and are distributed to the cortex chiefly of the flocculonodular lobe of the cerebellum. From here impulses pass to the red nucleus, some back to the vestibular nuclei, and to various other destinations.

(iv.) The vestibulo-reticular tract to the bulbar nuclei.

The Functions of the Labyrinth. The functions of the labyrinth may be divided into those of the vestibule, and those of the semicircular canals. They can be studied by the methods of extirpation, and stimulation. Both procedures result in disturbances of posture or of voluntary movements.

Unilateral labyrinthectomy in the rabbit leads to skew deviation of the eyes, nystagmus, rotation and lateral flexion of the head and spiral rotation of the neck and trunk towards the side of the lesion, diminution of tone of the limbs of the same side and increase in those of the opposite side, and frequently rolling movements, in which the animal turns over and over. In cats and dogs the deviation of eyes is lateral. The nystagmus is transitory, with the quick component to the normal side. The deviation of the eyes becomes maximal if the head is passively rotated to the normal side, but the nystagmus stops, and the rotation of the trunk ceases. The head and eye movements are direct effects of the extirpation of the labyrinth, while the neck and trunk reflexes are secondary and determined by the abnormal position of the head.

The rolling movements are only seen for a short time after the operation. Since the muscular tonus is diminished on the side of the lesion, the effect of any attempted progression movement is to cause the animal to turn over towards the injured side.

Bilateral Labyrinth Extirpation. The higher an animal is in the scale, the less important and the more readily dispensed with are its labyrinthine postural reflexes. Experiments on apes and cats show that, provided use of the eyes is retained, the animal behaves much as a normal one, except under two conditions.

The first of these is in falling free through the air. It is well known that a cat always falls on its feet, and this is accomplished by the employment of ordinary postural reflexes, especially labyrinthine ones. During the fall, first the head is righted, then the fore limbs and thorax, finally the hinder part of the body, aided by the tail. After bilateral labyrinthectomy this power of righting the body during a fall is lost.

The other case in which a response is lost is when the animal is placed in water. Under these circumstances the labyrinthectomised animal fails to orient itself, and would soon drown. The tonus of the muscles is not permanently lowered by removal of both labyrinths, but is kept up, as Sherrington has shown, by proprioceptive impulses from the muscles themselves, or from the muscles or other structures, such as joints, in other parts of the body, *e.g.* the neck.

The most usually accepted hypothesis is that the macular (otolith) organs are concerned chiefly with the *positions* of the head, and with *linear acceleration*, while the ampullæ of the semi-circular canals are concerned in notifying *movements*, particularly *rotations*. Whereas the former are largely static organs, the latter are entirely kinetic.

Two views have been held regarding the causes of stimulation of the maculæ: (*a*) That bending of the hairs causes stimulation; (*b*) that pulling on the hairs is the stimulus. As we shall see later, it is probable that the maximal stimulation is produced when the otolith organ hangs from and pulls upon, the hairs (A, Fig. 182), and the weakest stimulus when it pushes down vertically upon them (C, Fig. 182).

Denervation of the utricle was found by Tait and McNally to cause loss of muscular tone on the same side of the body.

The functions of the macula of the saccule are not clear. Denervation seems to have no evident results. It is possibly an organ that responds to vibration.

The responses of the labyrinth to various stimuli have been investigated by Ross,* who used the severed head of the frog, and led off the auditory nerve, and by Adrian,† using the cat, and leading off from the region of the vestibular nucleus. Responses were shown by action potentials in the nerve, and occurred when the preparation was subjected to tilting, angular acceleration, or deceleration or (in the frog), to vibration. The response to tilting or to linear acceleration was associated with the utricle macula, that to rotation with one of the canals and that to vibration, in frog, chiefly with the saccule macula. No other function has been attributed to the saccule.

Function of Semicircular Canals. It is now considered that each canal operates in one direction of rotation only. Thus, on rotating the head about a vertical axis so that the nose moves to the right, there is stimulation of one horizontal canal (the right). By rotation in the transverse plane, the right posterior canal ampulla is stimulated when the head moves to approximate right ear to right shoulder. Probably it is the bending of the hairs which is the essential factor. If the head is suddenly rotated about a vertical axis, the walls of the horizontal canals will rotate in the same direction. But the contained fluid, being rather viscous, and remaining stationary owing to inertia, will tend to apply forces to all structures with which it is in contact. The cupola and hairs in the ampullæ will then have a force applied to them tending to bend them, like reeds in a flowing stream. It has been objected that the canals have bores far too fine for any streaming of fluid to take place, but direct observation ‡ of the canals during rotation shows that movement does occur. The movement is slow,§ and if rotation is maintained, the cupola by its elasticity returns to its original position in the ampulla. On cessation of rotation the cupola is bent in the opposite direction.

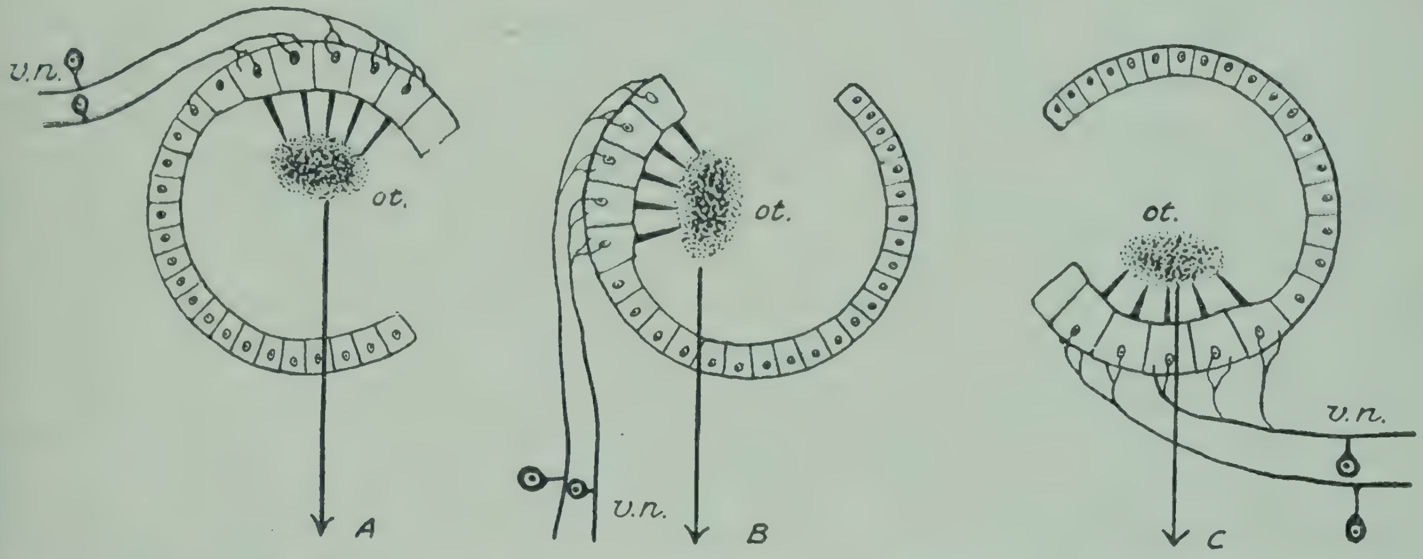


FIG. 182. Diagram to Illustrate the Influence of Position of the Head in Space on the Otolith Organ. In *A*, the otolith organ hangs vertically from the hairs, in *B* it pulls sideways, and in *C* it presses down on them, *v.n.*, fibres of the vestibular nerve. *ot.*, the otolith organ. The arrows indicate the pull of gravity.

Tait and McNally ‖ denervated one or more of the semicircular canals in the frog. When all the canals were denervated the animals assumed a normal posture, provided they were unmoved. Sudden tilting of the animal, however, failed to produce the normal compensatory movements of head and limbs. Their experiments show clearly that the horizontal canals only respond to rotations in the horizontal plane, while the vertical canals respond to rotation, each in its own plane. They infer that each horizontal canal is stimulated when moved in a direction away from its ampulla ("ampulla-trailing"), *e.g.* the right one only by turning the head to the right. The vertical ones, on the other hand, they believe to be stimulated when movement is such that the ampulla is foremost ("ampulla-leading"), *e.g.* the left anterior vertical by a movement of the head obliquely forwards and to the left (cf. Fig. 179). These results have been confirmed by Ross's direct recording of action potentials in the vestibular nerve, and by the work of Adrian.

---

* ROSS. *J. Physiol.*, 1936, **86**, 117.

† ADRIAN. *J. Physiol.*, 1943, **101**, 389.

‡ STEINHAUSEN. *Arch. f. d. ges. Physiol.*, 1933, **232**, 500.

§ LÖWENSTEIN and SAND. *Proc. R. S.*, 1940, **129B**, 256.

‖ TAIT and MCNALLY. *Amer. J. Physiol.*, 1925, **75**, 140, 155; *Q. J. Exp. Physiol.*, 1933, **23**, 147; *Phil. Trans.*, 1934, **224B**, 241.

Tonic movements of the musculature are produced when the cristæ of the canals are stimulated by suitable means, *e.g.* by rotations. The horizontal canals affect the musculature on the two sides of the body in opposite ways. When, for example, the whole animal is rotated to the right, the head itself turns to the left, the (extensor) tonus of the limbs on the right side is increased and the tonus on the left side is reduced. The vertical canals each affect a corresponding quarter of the body; thus stimulation of the left anterior vertical canal causes increased tonus of the left fore-limb, that of the left posterior canal increased tonus of the left hind limb, and so on.

Stimulation of the Semicircular Canals in Man. This can be carried out by rotation, or by galvanic or caloric stimulation. The results of rotating a subject on a turn-table or rotating chair at about one revolution per second, are similar to those obtained with animals. They comprise nystagmus, vertigo, and post-rotatory reactions. Nystagmus is a compensatory movement of the head and/or eyes, and consists of a slow reflex component, and a jerky (cerebral) return. They are due to the fact that the gaze tends to remain fixed on some stationary point. Hence, if rotation is to right, the eyes (and sometimes the head), in keeping the gaze fixed, move to the left slowly, and when the limit has been reached, jerk rapidly back again to the right to recommence the next slow movement. This is called a right horizontal nystagmus, the indication being conventionally given to the *quick* component. Vertical nystagmus can also be produced by appropriate movements. After the rotation, a nystagmus in the opposite direction occurs, due to inertia of the endolymph.

Vertigo, or giddiness is well known as a result of rotation. It is most marked after the rotation is over, and leads to compensatory movements in the opposite direction. Vertigo is a common clinical accompaniment of labyrinthine disorders. As a consequence of the tendency to fall over, various compensatory changes of muscular tone also occur after rotation, but are better seen after caloric stimulation.

Caloric stimulation (Barány's Test) is carried out by bending the head backwards at 60°, so as to bring the horizontal canals into a vertical plane, and then syringing the ear with cold water (20° C.). Vertigo and nausea result, and also definite after-effects, *e.g.* nystagmus to the opposite side, a curious twisting of the trunk, with extension of the arms, both of which are turned towards the side stimulated by cold ("discobolus position"); later the position is reversed. Another effect of the altered tonus is the "Past-pointing Test"—on attempting to touch a given point with eyes closed, the arm is deviated towards the stimulated side. The explanation of these caloric phenomena may be that convection currents are caused in the endolymph.

For Reference

Camis (1930). The Physiology of the Vestibular Apparatus. (Oxford: Clarendon Press.)

Jones (1918). Equilibrium and Vertigo. (Philadelphia: Lippincott.)

McNally and Stuart (1942). Physiology of the Labyrinth, etc. *War Medicine*, **2**, 683

MAXWELL (1932). Labyrinth and Equilibrium. (Philadelphia: Lippincott.)
QUIX (1925). The Function of the Vestibular Organ. *J. Laryng. Otol.*, **40**, 425, 493.

DECEREBRATE RIGIDITY

The state of decerebrate rigidity has already been mentioned as a release phenomenon, which is a caricature of the standing posture. The rigidity does not follow section of the medulla below the upper limit of Deiters' nucleus; thus, Section VI. of Fig. 183 does not produce it, but Section V. does; hemisection causes rigidity on the same side only. The rigidity is at once diminished or abolished if the vestibular nuclei (including Deiters') are destroyed,* or if the ventro-lateral columns of the cord are divided. If the vestibular nuclei on one side are destroyed, the rigidity on that side is abolished. These facts indicate the vestibular nuclei, and the vestibulo-spinal tract which descends from them, as essential for the condition, and suggest that these nuclei normally exercise on the extensor muscles a tonic influence, which is here exaggerated by release from the modifying influence of some centres still higher up.

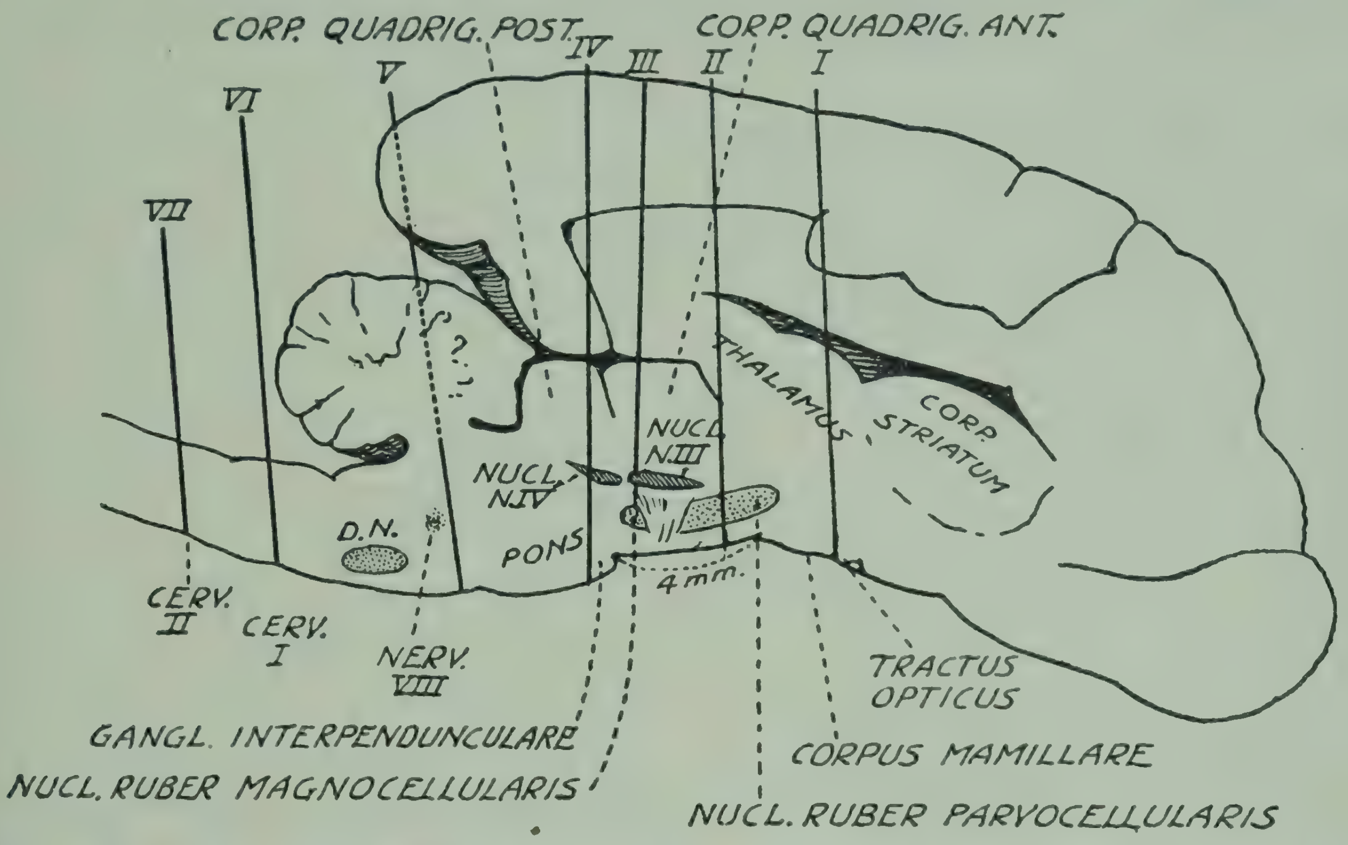


FIG. 183. Longitudinal Section of Cat's Brain, showing the various sections of the brain stem in their relation to decerebrate rigidity.

Section I. The Thalamus animal. Possesses heat regulation and normal postural reactions.

Section II. Mid-brain animal. Temperature regulation lost; posture normal.

Section III, below red nucleus. Decerebrate rigidity.

Section IV, decerebrate rigidity present.

Section V, just above vestibular nuclei; decerebrate rigidity.

Section VI, below vestibular nuclei (D.N.); no rigidity. Neck reflexes still present.

Section VII, spinal animal. No neck reflexes. (Re-drawn, after MAGNUS.)

The location of these higher centres is ascertained by making sections from before backwards. It will be seen from Fig. 183 that rigidity does not then appear until the section passes through the mid-brain, just below the red nucleus, or through its anterior part. It is inferred, therefore, that centres which in the normal condition keep the vestibular nuclei from overacting are at the level of the red nuclei, but it does not necessarily follow that the red nucleus is the only centre concerned. Doubtless it is one of them, but destruction of the red nucleus alone does not cause rigidity.† Nor is it certain that section of the rubro-spinal tract alone will cause rigidity.

---

* FULTON, LIDDELL and RIDDOCH. *Brain*, 1930, **53**, 327.

† INGRAM and RANSON. *Amer. J. Physiol.*, 1932, **102**, 466.

The rigidity is of reflex origin, the afferent impressions concerned arising from the muscles themselves, and from the labyrinth, and being in both cases relayed from the vestibular nuclei. The afferents concerned in the rigidity of the hind limbs come from the muscles only; hence, division of the afferent roots of the muscular nerves abolishes it, while removal of the skin does not. For the fore limbs, de-afferentation does not remove the rigidity, but destruction of the labyrinths as well will do so.*

These relations are shown diagrammatically in Fig. 184.

The muscles which become tonic in decerebrate rigidity are those which counteract the effects of gravity, in maintaining the normal posture of the animal. These are, in most animals, the extensor muscles. In the sloth, which hangs upside down, it is the flexor muscles which are the tonic ones.

Decerebrate Rigidity in Man. A condition closely resembling decerebrate rigidity is also seen in the human subject as a result of certain lesions of the brain stem, *e.g.* the result of suprapituitary tumours. The subject shows extension of the legs, with the feet pointed down in plantar flexion; the arms are semi-flexed at elbows, forearm pronated, and wrists and fingers flexed † (Fig. 185).

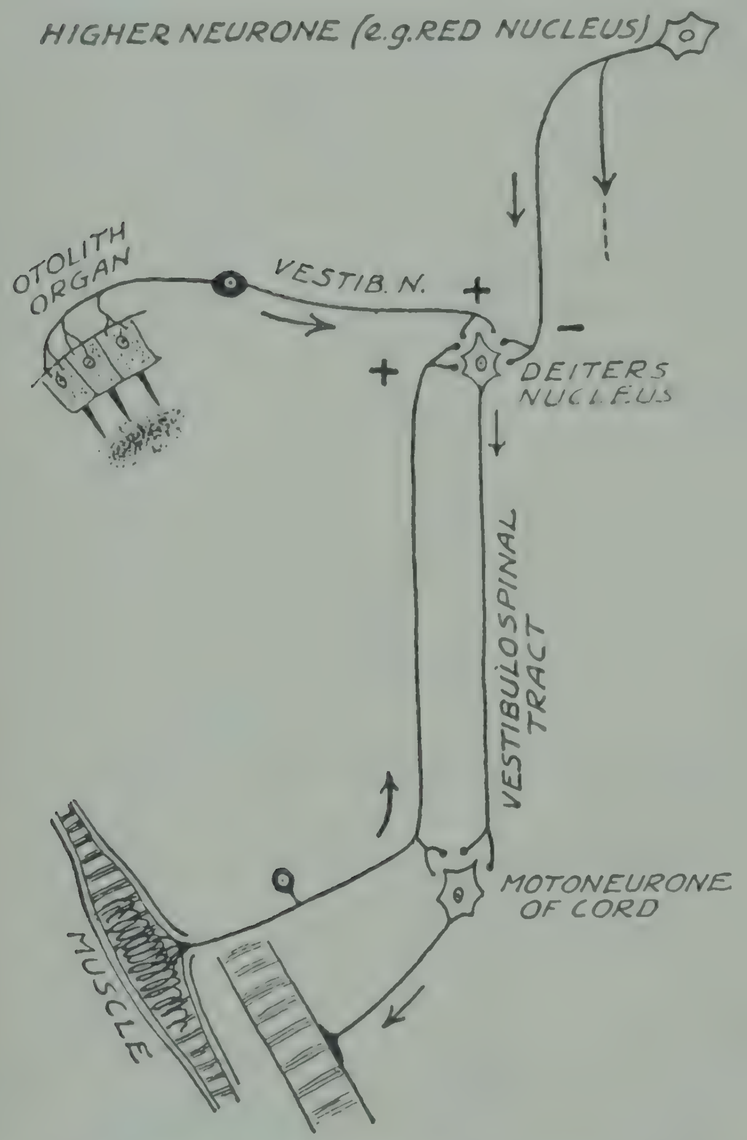


FIG. 184. Diagram to Illustrate the paths concerned in Postural Tonus and Decerebrate Rigidity (Interneurone in cord omitted.)

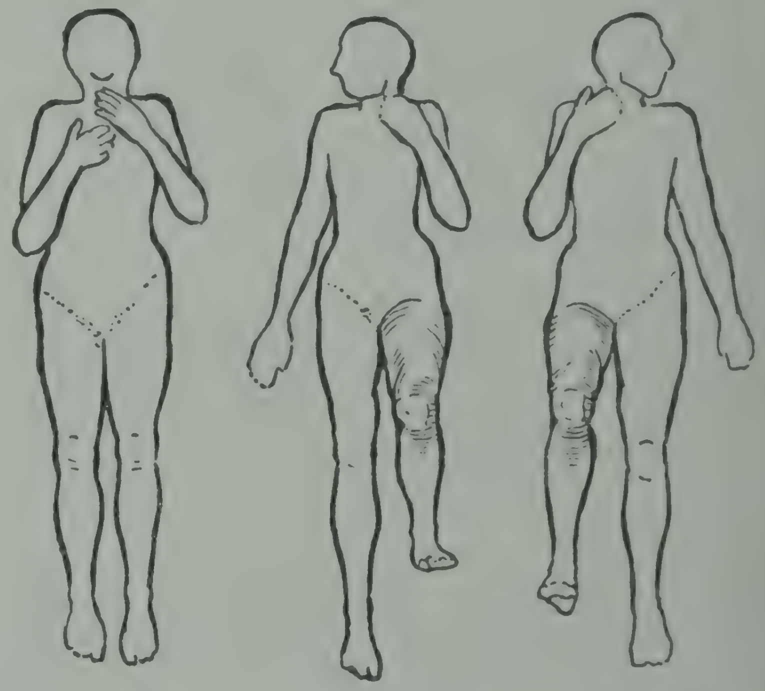


FIG. 185. Decerebrate Rigidity in Man. (SAMSON WRIGHT, after WALSHE, "Applied Physiology," Oxford University Press.)

Rotation of the head causes extension of the arm on the side to which the nose is turned, and flexion, with augmented Babinski response, of the opposite leg (Fig. 185). This is a postural reflex, which is explained later.

THE POSTURAL REFLEXES

The higher animals normally exhibit characteristic attitudes in repose or action, by which they are oriented with reference to gravity. To any circumstance which tends to displace the animal from its posture there is offered a resistance which prevents the displacement, or which tends to restore the animal to its correct attitude.

---

* POLLOCK and DAVIS. *Amer. J. Physiol.*, 1931, **98**, 47.

† WALSHE. *Lancet*, 1932, **101** (ii), 644.

Static postures are often maintained for long periods of time without fatigue. The state is one of tonic contraction of the muscles concerned, and the assumption of any posture is the result of a delicate condition of balance in the muscular system. In the maintenance or recovery of posture many separate types of reflex converge to produce a common result; these reflexes arise from the labyrinth, the eyes, and the muscles themselves.

The removal of the cerebral hemispheres has but little effect on normal posture, or on the righting reactions by which that posture is resumed after having been interfered with. It is clear, therefore, that these reactions are true reflexes. In a fully conscious animal, of course, voluntary movements may be added to the reflex reaction.

Some muscular tonus persists in the spinal animal (Fig. 186B), but this is slight and bears no resemblance to normal posture. The bulbo-spinal animal, by contrast, shows the condition of *decerebrate rigidity* (Fig. 186A), which is due to an exaggerated state of tonus of those muscles (extensors) which are employed in the posture of standing. Such a decerebrated animal can, in fact, maintain a standing posture, but when upset it makes no movements which would tend to place it in its former standing pose. It was shown by Sherrington that the tonus distribution in the decerebrate preparation is not entirely proprioceptive, but depends, for instance, on the position of the head.

When the whole of the mid-brain is left as well, the attitude of standing is normal (Fig. 186C), and, moreover, the animal will now of its own accord return to the standing posture when displaced from it. Most of the postural reflexes of the "thalamus" animal (*i.e.* with intact thalamus and mid-brain) are practically identical with those of the normal animal. (The advantage of leaving the thalamus is that temperature regulation is retained.) This fact, and the knowledge which we derive from our own experience, teach us that the afferent impulses upon which posture is based and adjusted but rarely obtrude into the field of consciousness, and can, therefore, only be studied by objective methods. Such studies have been made by Magnus, and to translate his own words freely, have shown "that many different sense organs are involved, that very different centres integrate their stimuli, that the muscles are concerned in many different combinations in the production of the reactions, and that a whole series of diverse reflex groups are balanced together."

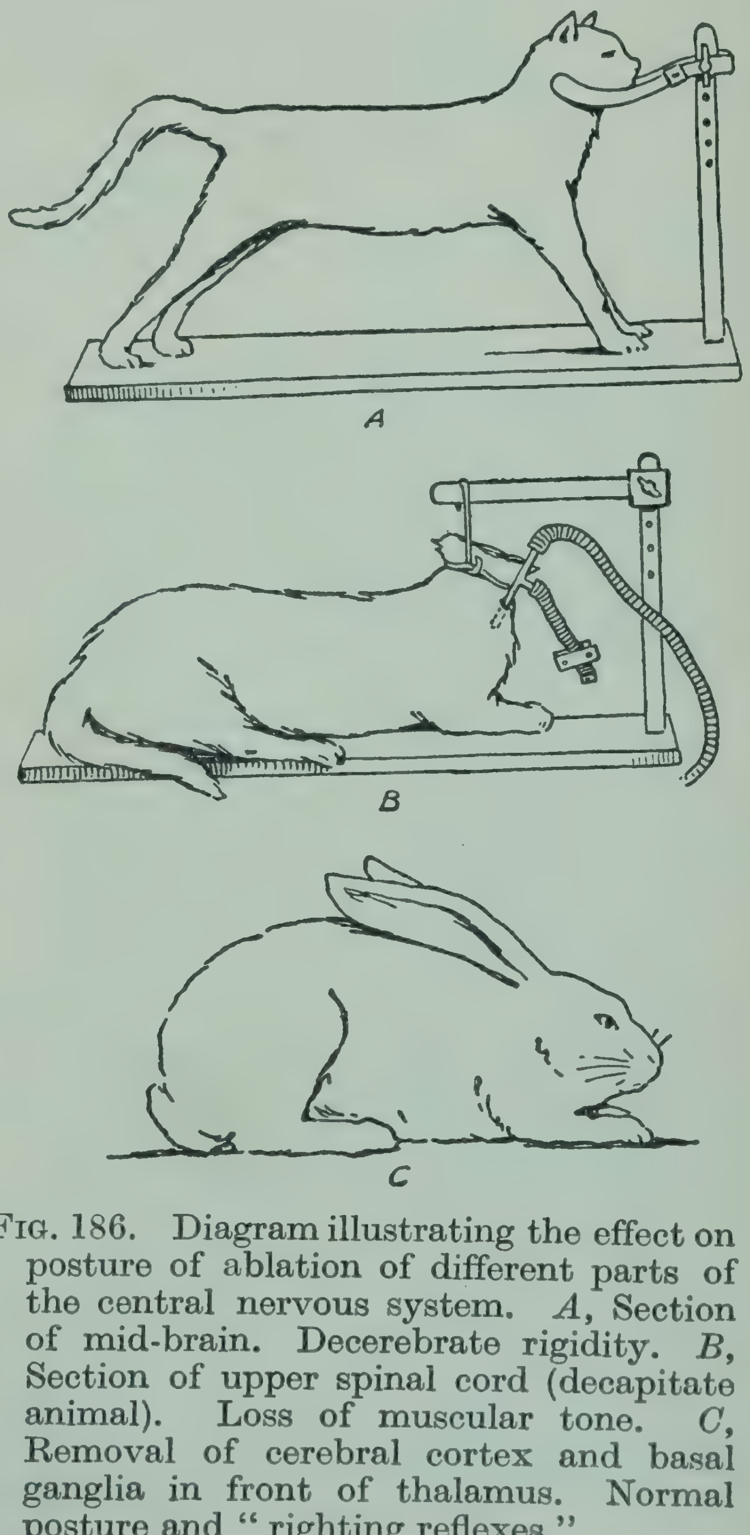


FIG. 186. Diagram illustrating the effect on posture of ablation of different parts of the central nervous system. *A*, Section of mid-brain. Decerebrate rigidity. *B*, Section of upper spinal cord (decapitate animal). Loss of muscular tone. *C*, Removal of cerebral cortex and basal ganglia in front of thalamus. Normal posture and "righting reflexes."

The postural reactions of the body when at rest are called *static reflexes*, and fall into two groups: the *stance reflexes* or reflexes of *pose*, in which the body, or a part, remains in a definite attitude, and *righting reflexes*, by means of which the pose is restored after disturbance. The former, so to speak, "keep you right way up when you are" while the latter "put you right way up when you are not."

The postural reactions of the body when in movement are called *stato-kinetic reflexes*, and are produced in a definitely predictable manner by the movement itself, which in turn is largely guided by them under normal conditions of life.

These important postural reflexes fall into the following heads and it will be noted that there is a good deal of inter-dependence :—

I. STATIC REFLEXES

(i) *Local Static Reactions*, *e.g.* confined to one limb.
(ii) *Segmental Static Reactions*, involving one segment, *e.g.* both hind limbs, both fore limbs, or the neck.
(iii) *General Static Reactions*, involving more than one segment. These may be again divided into :—

A. *Reflex Stance.*
  (1) Influence of the Head on Stance.
    (*a*) Effect of Tonic Neck Reflexes on Limbs.
    (*b*) Effect of Tonic Labyrinth Reflexes.
      (α) On the Limbs.
      (β) On Neck and Trunk.
    (*c*) Collaboration of Neck and Labyrinth Reflexes.
    (*d*) Indirect Influence of Labyrinth on Limbs through Action on Neck Reflexes.
  (2) Indirect Influences on Posture.

B. *Compensatory Pose of Eyes.*
  (1) Effect of Tonic Labyrinth Reflexes on Eyes.
  (2) Effect of Tonic Neck Reflexes on Eyes.
  (3) Collaboration of Tonic Labyrinth and Neck Reflexes on Eyes.
  (4) Collaboration of Compensatory Eye Poses.

C. *Righting Reflexes.*
  (1) Effect of Labyrinth Reflexes on the Head.
  (2) Effect of Body Posture on the Head.
  (3) Neck Postural Reflexes.
  (4) Effect of Body Postural Reflexes on the Body itself.
  (5) Visual Postural Reflexes.

II. STATO-KINETIC REFLEXES

A. *Reactions to Rotation.*
  (1) Head Reactions.
  (2) Eye Reactions.

B. *Reactions to Progressive Movement.*
  (1) Reactions of Head.
  (2) Reactions of Extremities.

We must now consider these types.

I. Static Reflexes

(i) LOCAL STATIC REACTIONS

It is a remarkable fact that a limb which at one time is freely movable, at another time is converted into a rigid pillar capable of supporting the weight of the body. The formation of the rigid support by simultaneous contraction of flexors and extensors is due to a local reflex, called by Magnus the *positive supporting reaction*, while its resolution, enabling the limb to be used for other purposes, is also a local reflex, called the *negative supporting reaction*. These reflexes receive their explanation from the researches of Sherrington and Liddell, and of Magnus and his collaborators, Rademaker, Schoen and Blake Pritchard.

The afferent stimuli for the positive supporting reflex are in part proprioceptive stimuli evoked by movements of the distal parts of the limbs, by which some of the flexor muscles are stretched, and in part are due to exteroceptive stimuli from the skin, *e.g.* contact of pad of foot with the ground. A direct consequence of the stretching is that these flexor muscles themselves are reflexly contracted (stretch reflex); a second result is that there is reflex contraction of flexors and extensors, whether stretched or not, the antagonistic muscles thus acting synergically in fixation, at the proper angle, of the joints concerned. In this complex local reaction very different reflexes co-operate, reinforce, and if necessary replace each other, so that standing under very different positions of the joints is possible.

In the negative supporting reaction, the active stimulus is proprioceptive, due to volar or plantar flexion which unlocks the series of positive reflexes and loosens the limb for action.

(ii) Segmental Static Reactions

A good example of these is the crossed extension reflex already studied.

(iii) General Static Reactions

A. Reflex Stance. (1) *Influence of the Position of the Head.* Experiments on the effect of the position of the head on the tonic pose of the body are carried out on decerebrate animals. The distribution of tonus is modified as a result of stimuli from the neck when the head is moved with reference to the trunk, and also, by the operation of labyrinthine reactions, when the head (with the body) has its position altered in space. These two effects can be separated experimentally.

(*a*) *Tonic Neck Reflexes.* These are obtained in a pure form on animals deprived of both labyrinths, and are also seen in certain clinical cases. When the head is moved by bending or twisting the neck, the extensor tonus of the decerebrate preparation shows modifications which persist so long as the altered relation of head to trunk is kept up. These alterations are due to altered tensions on the neck muscles when the head is moved. The muscles of the right and left sides or of the fore and hind limbs behave in opposite ways.

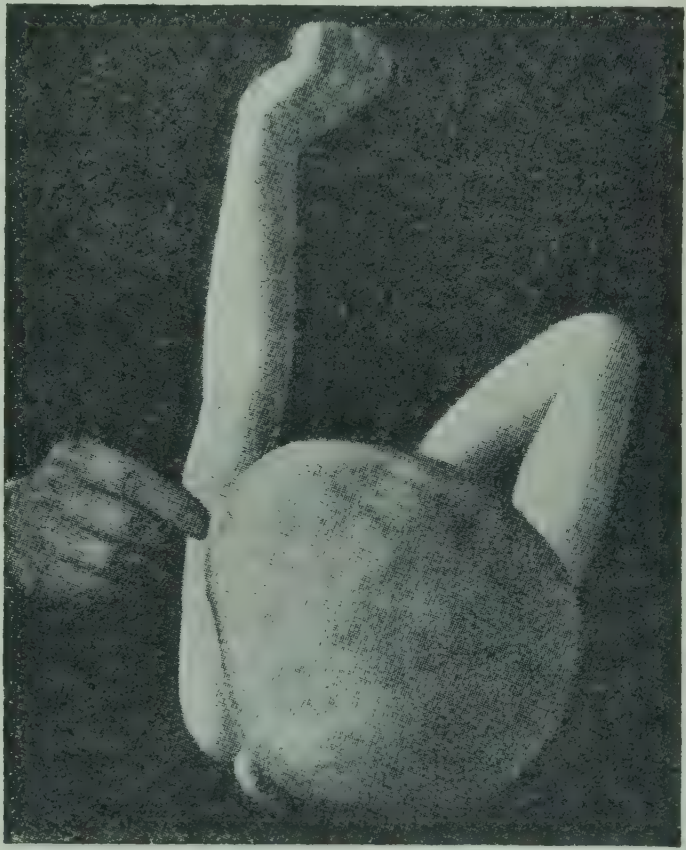


Fig. 187. Hydrocephalic infant. Turning of head to left gives increased tonus of left arm, due to tonic neck reflex. (Magnus, "Körperstellung.")

When the head is rotated on the atlanto-epistropheal joint, the extensor tonus of the fore and hind limbs is increased on the side towards which the nose is turned, and diminished on the other side (Fig. 187).

When the head is inclined sideways by bending the neck, so that, say, the left ear approaches the left shoulder, the extensor tonus of the limbs is increased on that side, while on the right side there is a diminution. This reaction to bending of the neck obviously serves the purpose of keeping the centre of gravity of the body over the feet.

Bending the neck backwards causes increased tonus of the fore limbs, with reduction of tone of the hind limbs. This results in a posture such as is seen in a normal cat craning the neck upwards to look on a table. When the head is bent downwards, the tonus distribution is in the opposite direction, as in a normal cat eating food.

The centre which is responsible for these neck reactions is located in the upper cervical spinal cord and lower medulla.

(*b*) *Tonic Labyrinth Reflexes.* These reflexes are best studied on decerebrate cats with intact labyrinths, and fixed to a holder in such a manner that the head cannot be moved on the trunk.

The most conspicuous effect is that on the limbs, all of which react similarly. Maximal extensor limb tonus is seen when the animal is placed on its back, so that the mouth cleft is inclined upwards at 45 degrees to the horizontal; when the animal is placed on its belly so that the mouth cleft is downwards at 45 degrees, the extensor tonus falls to a minimum. At intermediate positions the tonus is of intermediate intensity.

The receptors concerned in these reflexes are the maculæ, chiefly those of the utricles. The utricle macula lies in a horizontal position when the head is normally oriented, and the otolith then lies above the hair cells which spring from the macula. The macula of the saccule is situated on the medial wall, and is, therefore, vertical when the head is normally placed, the hairs of its hair cells lying horizontally in the otolith. The exact orientation of these maculæ in the rabbit is illustrated in Fig. 188, which shows a rabbit's skull in the normal position in space, with enlarged models of the maculæ, represented by oval plates, in their proper spatial arrangement, attached to it. The surface of the macula to which the otolith is attached is indicated in the utricle by white dots, and in the saccule by white stripes.

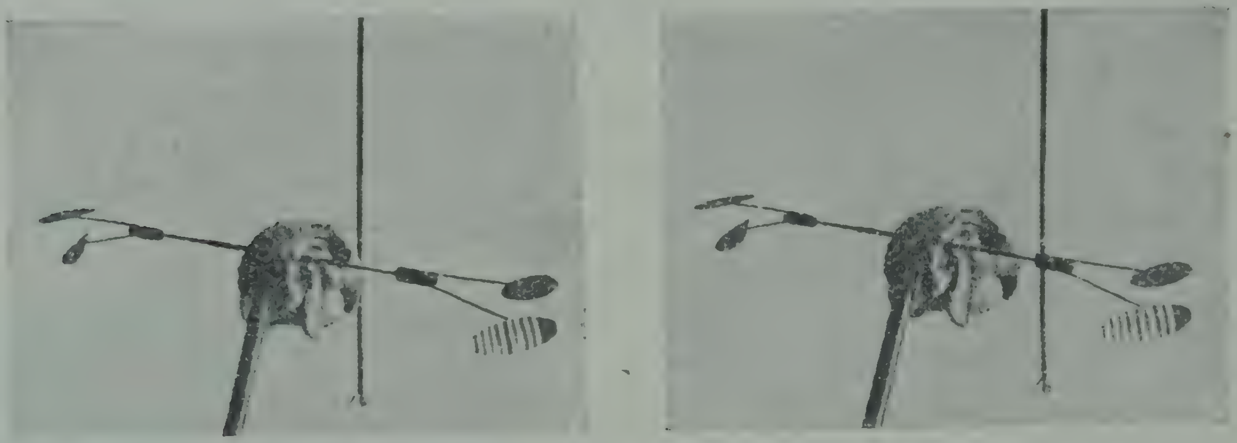


FIG. 188. Rabbit's Skull with Oriented Magnified Models of the Utricle and Saccule Maculæ. Skull in normal position. Stereoscopic. (MAGNUS, "Körperstellung.")

When the animal is placed in the position which gives the maximal labyrinth tonus it is found that the utricle otolith is hanging vertically beneath its macula; when the head is placed in the position which gives the minimum of limb tonus, then it is found that the otolith is vertically over its macula. Hence it is concluded that the otolith produces the greatest stimulus when it hangs from, and so pulls upon, its hair cells, and the least when it presses them downwards on to the macula. The stimuli are continual, and give rise to steady reflex tonus so long as the position is kept up. Similarly, the saccule macula is most stimulated when the otolith hangs from it, as on the left side when the left side of the head is below. Hence there arise, from the asymmetric nature of the stimuli from the two saccules, asymmetric states of tonus, which constitute the righting reflexes of labyrinthine origin.

Deprivation of both labyrinths leads to complete disappearance of the reactions to altered orientation of the head in space.

The tonic neck and labyrinth reflexes show an algebraic summation in the normal animal, in which they both have full play. It is also found that the labyrinth exerts an indirect action by giving rise first to tonus changes in the neck, which produce neck movements, and hence, by the operation of the neck reflexes, exert an effect on the limbs. The two effects of the labyrinth on the body are algebraically summated.

In these postural reflexes the red muscles play a larger part than the pale muscles (Fig. 189).

(2) *Indirect Influences on Posture.* Posture in the normal animal is largely influenced by impressions received from the outer world through the senses of sight and hearing. These lead to voluntary movements of the head and, by the operation of the reflexes so far described, to an automatic modification of the pose. To take the cat as an example, we have already referred to certain characteristic postures related to neck reflexes, and a little consideration will show that the pose will be further reinforced by the action of the labyrinth reflexes.

Suppose that a cat hears a mouse moving on its right. The head is turned to the right; this alters the centre of gravity, but, as a result of the tonic neck reflex, there is an increase of tonus on the muscles on the right side of the body, which preserves the balance by throwing the weight on the right limbs. Now, because the left limbs are less loaded, they will be the first to move if the animal springs to its prey, so that the cat will automatically move in the right direction.

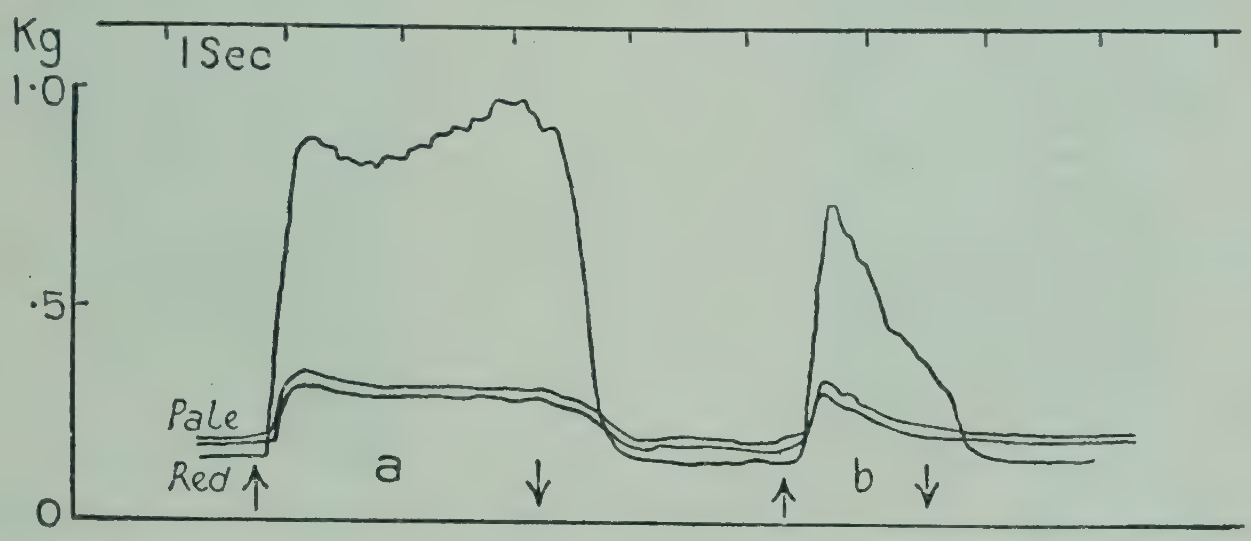


FIG. 189. Response of Pale (Short Lat. Head Triceps) and Red (Short Med. Head Triceps) Muscle to Labyrinthine and Neck Reflexes. Mid-brain preparation. ↑ dorsiflexion and ↓ ventriflexion of neck. Tracing from pale muscle shown by double line and from red muscle by single line. (CREED, *et al.* "Reflex Activity of the Spinal Cord." Oxford Univ. Press, 1932.)

B. Compensatory Pose of the Eyes. The compensatory pose exhibited by the eyes under certain conditions is really a special case of stance reflexes. In man and other animals with frontal vision it is of much less importance than in such animals as the rabbit, which have lateral vision. Like other muscles, the tonus of the eye muscles may be modified as a result of reflexes arising either from the labyrinth or from the neck. The outcome is that the eyes take up a position such that their visual fields suffer as little change as possible.

(1) *Tonic Labyrinth Reflexes on the Eyes.* In these reflexes no effect is produced by way of the labyrinth unless movements alter the relation of the head to the horizontal plane. The responses are to rotation of the head (and body) about the occipito-nasal axis (" vertical " response) and rotation around the bi-temporal axis (" rotatory " response). In the performance of these movements there is always a delicately adjusted reciprocal action of antagonistic muscles in such a manner that the eyes continue, as it were, to look at the same object in spite of the movements of the animal. Thus if the animal in the prone position is tilted so that the head moves downwards, the eyes will move upwards, and *vice versâ*.

The centres for all labyrinth reflexes on the eyes lie between the vestibular nerve entry and the nuclei for the eye muscles.

(2) *Tonic Neck Reflexes on the Eyes.* In these reactions, which are obtained after destruction of the labyrinth, when the head is moved with relation to the body, *all* the eye muscles are brought into play, and compensatory eye poses are obtained in response to alteration of the head with reference to any plane, including the horizontal. The effects are similar to those obtained with the labyrinth reflexes.

Under normal conditions the labyrinth and the neck reflexes are both present and are summated algebraically, so that, while neither reflex alone is enough to compensate fully for head movements, the addition of the two does so. The exception to this is when the movement of the head is in the horizontal plane, in which case, as already stated, there is no labyrinthine reflex set in operation.

C. Righting Reflexes. The reactions by which an animal keeps itself right side up and recovers its normal posture after it has been disturbed, are well illustrated in the familiar example of the falling cat, which always lands on its feet. The analysis of these reflexes is carried out best in thalamus animals, so that voluntary movements are eliminated. An animal takes up by means of these reactions a normal stance, which is the basis for other voluntary or reflex movements. Righting reflexes are present at birth, and are carried out by co-operation of five different groups of reflexes. When, by suitable procedure all of these are eliminated, we have the "*zero condition*," in which no righting reflexes are shown. In rabbits this is achieved by destruction of both labyrinths, and holding the animals free in the air; cats, dogs and monkeys, which use the eyes for righting reactions must also be blindfolded if the cerebrum is intact. Knowing the zero conditions we can compare the effect of introducing the responsible reflexes one by one.

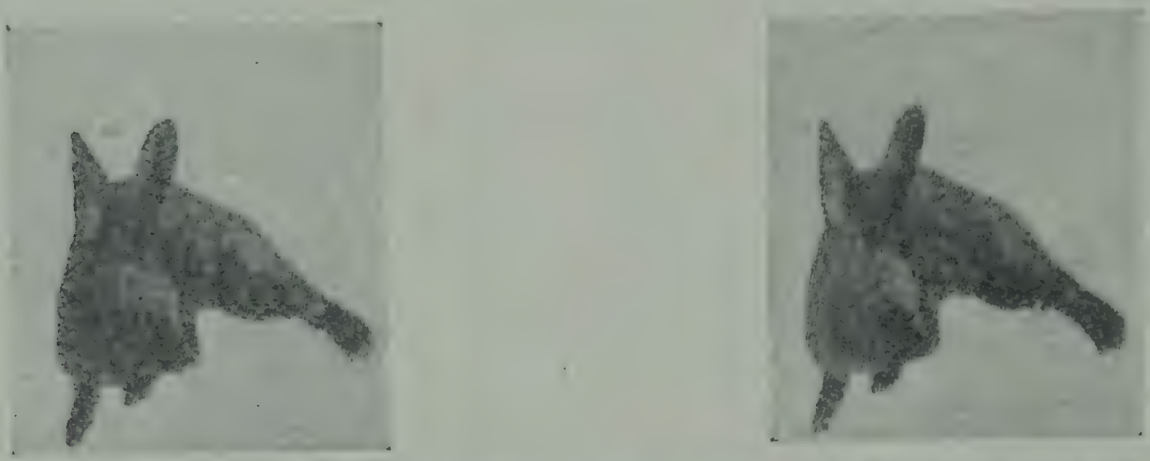


FIG. 190. Response of Thalamus Rabbit to Rotation of Pelvis when held Free in Air. The anterior part of the body retains the normal position owing to co-operation of labyrinth, and neck-righting reflexes. View with stereoscope. (MAGNUS, "Körperstellung.")

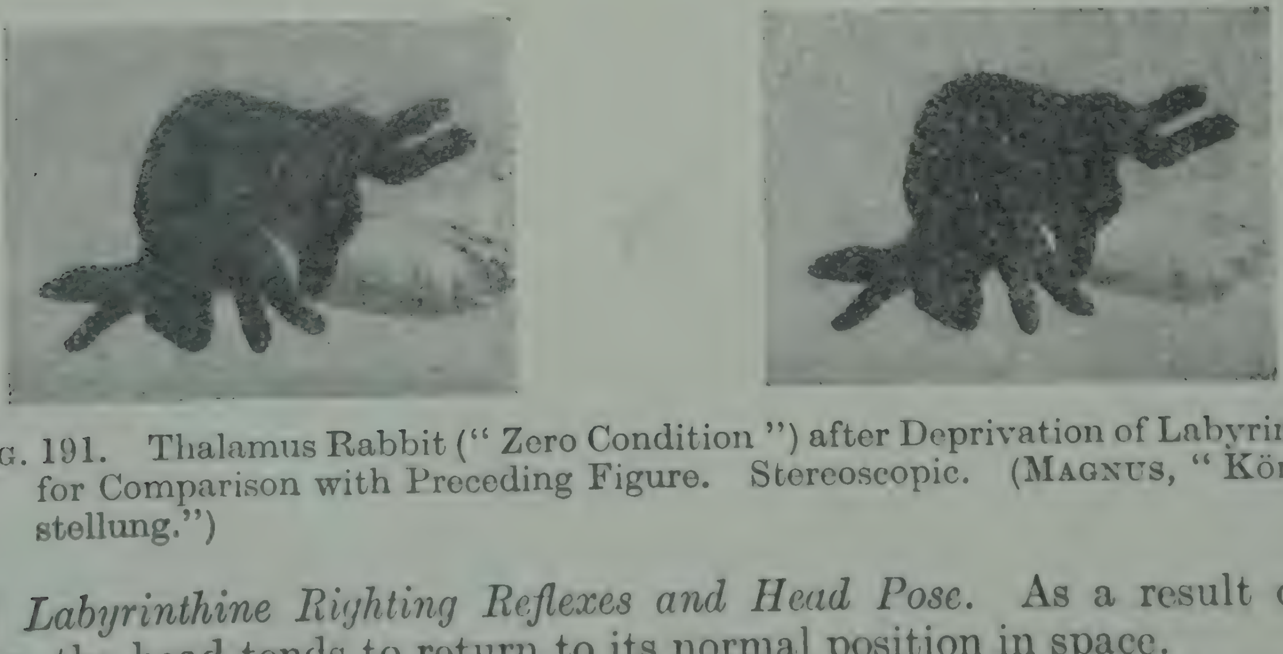


FIG. 191. Thalamus Rabbit ("Zero Condition") after Deprivation of Labyrinths, for Comparison with Preceding Figure. Stereoscopic. (MAGNUS, "Körperstellung.")

(1) *Labyrinthine Righting Reflexes and Head Pose.* As a result of these reflexes, the head tends to return to its normal position in space.

When a thalamus rabbit is held up in the air by gripping the pelvis (to avoid contact stimuli which would be introduced if the animal were allowed to lie on the ground), it is seen that the head tends to remain in the normal position when the pelvis is moved to the most various positions. This is accomplished by appropriate twisting or bending of the vertebral column (Fig. 190). After extirpation of both labyrinths the reflexes are abolished (Fig. 191).

(2) *Body Posture on the Head.* If a rabbit deprived of both labyrinths be laid on the ground, the head again tends to assume the normal pose. This is due to the operation of asymmetric stimuli arising from contact of only the lower side with the ground. Thus, if the animal is laid on its right side, contact stimuli will arise from that side, and not from the left, and these lead to orienting reactions of the head (Fig. 192). If the stimuli are made symmetrical, as by placing the animal in the lateral position and laying a board on its upper side, the head orientation vanishes (Fig. 193).

Owing to these two reflexes from labyrinth and from contact stimuli, there results a twisting of the neck when the animal is displaced, and this, in its turn, leads to the production of

FIG. 192. Thalamus Rabbit without Labyrinths laid on Right Side. The head, and then the anterior half of the body, assumes the normal pose owing to asymmetric contact stimuli. Stereoscopic. (MAGNUS, "Körperstellung.")

(3) *Neck Reflexes.* These operate in much the same way as the tonic neck reflexes, and the result of the combination of all these reflexes is that first the front and then the hind limbs follow the posture taken by the head, so that ultimately the whole animal comes to resume its normal pose. For instance, suppose an animal is laid on its left side. In consequence of labyrinthine and contact reactions, the head is turned towards the left shoulder (*i.e.* remains horizontal); this leads to neck reflexes and increase of tonus of the left fore limb, which rights the fore part of the body, then to increase of tonus in the left hind limb, by which the hind part of the body is also righted.

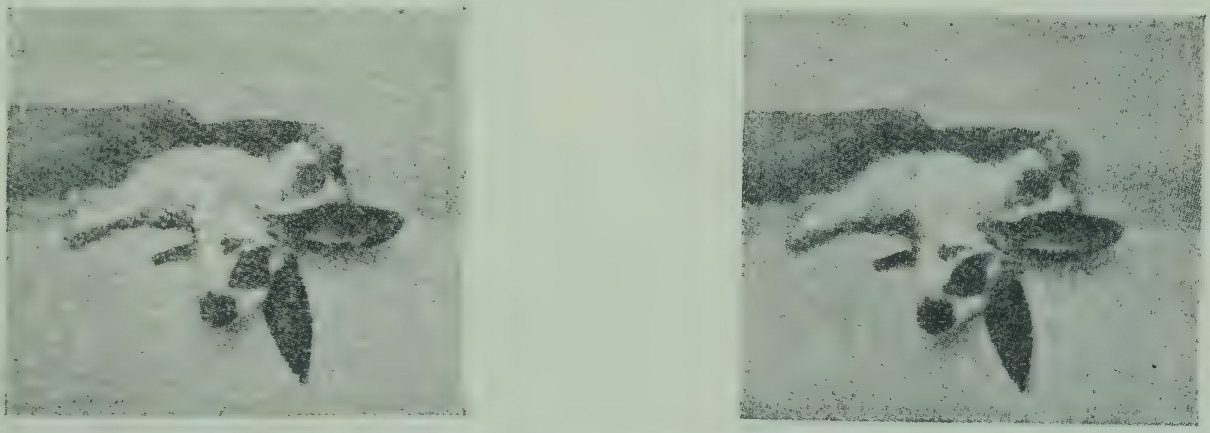


FIG. 193. The same Rabbit as last Figure laid on Side, but with Weighted Board placed on Upper Side; no Righting Reaction. Stereoscopic. (MAGNUS, "Körperstellung.")

(4) *Reflexes from the Trunk upon the Body itself.* The effect of asymmetric contact stimuli such as arise when an animal is laid in the lateral position affects not only the head, but also the body itself, so that if the head is fixed in the lateral position the body may be righted even though this entails opposition to the neck and labyrinth reflexes. In the normal animal there is thus a double reflex mechanism, which is summated to produce the total effect.

In decorticated primates lying on the side, the limbs on the under side are extended, and those on the uppermost side are flexed, and the digits in a grasping position—the "grasp-reflex."

(5) *Visual Righting Reflexes.* These reactions can be studied in cats and dogs with intact cerebral cortex after bilateral labyrinthectomy. The animals can orient their heads quite well if allowed to use the eyes, but when these are covered they revert to the disoriented position of the head, just as the labyrinthectomized rabbit does (Figs. 194 and 195).

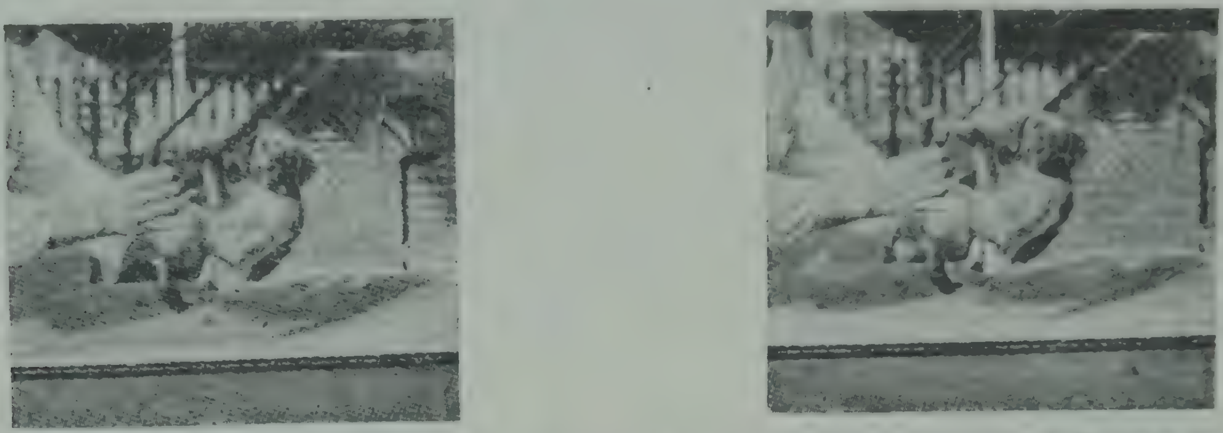


Fig. 194. Dog Sixteen Days after Double Labyrinthectomy, the Pelvis held in the Air in Horizontal Dorsal Position; Righting of Head and Thorax. Stereoscopic. (MAGNUS, "Körperstellung.")

Centres for Righting Reflexes. Visual righting reflexes are cortical. It has been shown by Rademaker, using the method of progressive section at various levels, that the other centres responsible for the righting reflexes are situated in the mid-brain and pons. The neck-righting reflexes are located in the pons, the remainder are in the ventral portion of the mid-brain at about the level of the third nerve nucleus, and probably in the large-celled portion of the red nucleus from which the rubrospinal tract arises. Destruction of the pyramidal tract, substantia nigra, or cerebellum leaves the righting reactions unchanged. Mesial sagittal section of the mid-brain causes disappearance of righting reflexes if Forel's decussation has been divided, not otherwise.

The spinal centres appear to be under the influence of higher centres of two types: medullary centres, which tend to cause over-action of the extensor muscles, and the red nucleus, which, acting *viâ* the rubrospinal tract, tends to counteract this by throwing the flexors into action. By the combined action of the two we get in the thalamus animal a normal distribution of tonus; the pyramidal tracts in the intact animal also exert a flexor effect on the spinal centres.

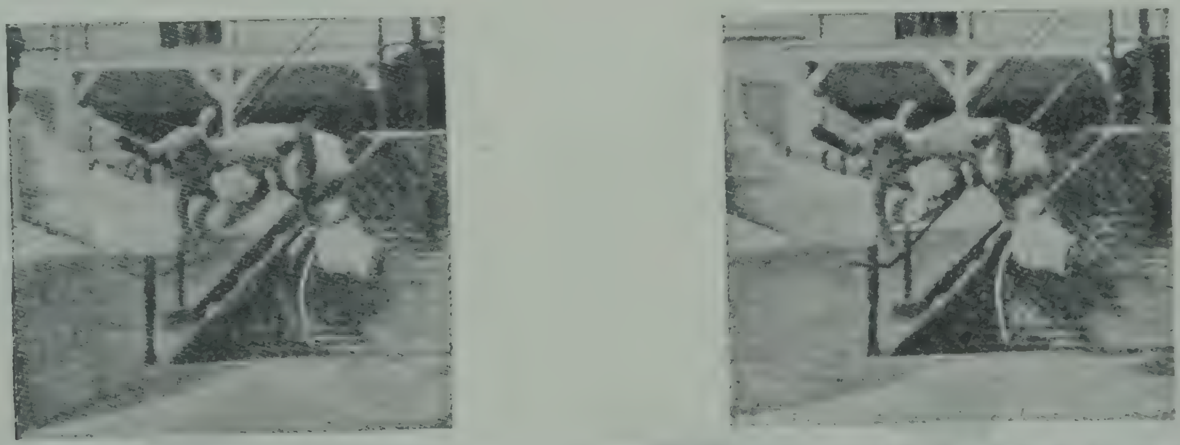


FIG. 195. The Same Dog held in the same Manner, but with the Eyes Covered; no Righting Reaction. Stereoscopic. (MAGNUS, "Körperstellung.")

II. Stato-Kinetic Reflexes

Unlike the righting reflexes, these are produced by the actual movement and not by any abnormal orientation. In so far as they arise from the labyrinth, they may be said to be due, not to movement in itself, but to an acceleration, positive or negative. This may be either angular acceleration, *i.e.* movements of rotation, or progressive acceleration when the movement is i a straight line.

A. Reactions to Rotation. (1) *Reactions of the Head.* When an animal or a man is rotated, say, to his right, the eyes tend to remain fixed on a given point, and hence the head is steadily moved to the left as far as possible, and then sharply to the right, this movement being repeated. This is called *head nystagmus*, with the rapid component in the same direction as that of the rotation. After the rotation has been stopped the nystagmus is in the opposite direction, *i.e.* in the case supposed, with the rapid component to the left.

All these phenomena may be explained as due to the effects of the rotation on the semicircular canals.

(2) *Eye Reactions.* The reactions of the eyes to rotation are similar to those of the head ; that is to say, eye nystagmus, in the same direction as the rotation, is produced. The effect on the eyes is naturally best seen when the head is fixed with relation to the body. On cessation of rotation the nystagmus is reversed.

When the head instead of the whole body is moved, the response of the eyes to the movement of rotation so produced leads to their moving in a direction towards that in which they finally become fixed by the static reflexes already considered.

B. Reactions to Progression Movements. (1). *On the Head.* These are best seen in the dog when stood on a board. When moved suddenly upwards, the head is moved ventrally and the fore-paws flexed ; when the movement ceases, the head moves dorsally, and the fore-paws are extended. On movement downwards the opposite effects are produced.

(2) *On the Extremities.* When the animal is held head downwards, the reaction to sudden downward movement is that the fore-limbs are moved in an oral direction, as in preparing the animal to take up its weight on falling to the ground.

This is called a " placing reaction." It is dependent on integrity of the cerebral cortex, and depends on impressions from eyes, labyrinth and proprioceptors.*

C. Reaction to Movements of Parts of Body. When a part of the body is moved, reflex adjustment of pose is often necessitated to maintain the balance, and the responses described above are variously brought into play. Thus, if a dog has an ulcerated foot, that leg is held in a position of flexion, and the posture and the movements of locomotion are arranged in accordance with the modified support which the body now has. This is all reflex, and depends on the constant arrival, at the centres concerned, of stimuli from the affected foot.

In general we can say that by means of the stato-kinetic reflexes the body is brought into its normal stance, which pose is maintained by the operation of the various static reflexes. It will have been noticed that the head plays an important part in deciding most of these responses. The head contains the organs of special sense, particularly the distance receptors for sight and sound, so that the head tends to be voluntarily directed towards stimuli which affect these sense organs. The proper orientation of the remainder of the body which ensues is effected by the action of the various reflexes which have been described.

For Reference

Hellebrandt and Franseen (1943). Physiological Study of the Vertical Stance in Man. *Physiol. Rev.*, **23**, 220.

Magnus (1924). Körperstellung. (Berlin : J. Springer.) (In which is given a full bibliography) ; Cameron Prize Lectures, *Lancet*, 1926, September 11th (p. 531), and September 18th (p. 585) ; Croonian Lecture. *Proc. Roy. Soc.*, 1925, **98B**, 339.

Langworthy (1928). The Control of Posture by the Central Nervous System. *Physiol. Rev.*, **8**, 151.

Maxwell (1932). Labyrinth and Equilibrium. (Philadelphia and London : Lippincott Co.)

Rademaker (1931). Das Stehen. (Berlin : Hirschwaldsche.)

---

* Bard. *Harvey Lect.*, 1937–38, **33**, 143.

CHAPTER XIX

HIGHER AFFERENT PATHS

Many of the afferent fibres of the spinal cord, in addition to furnishing collaterals which subserve the purposes of spinal reflexes, ultimately ascend to higher levels. Some, after various relays, ultimately reach the cerebrum, and convey impressions to consciousness ; others establish connections with higher reflex centres ; some do both. These various paths must now be briefly considered, together with those of the more specialised senses. Usually three neurones, primary, secondary and tertiary (or first, second and third order), are met with in the course of an afferent path from periphery to the highest centre, but auditory and visual paths have more.

I. CONSCIOUS SENSATIONS

So far as the brain stem is concerned, the main afferent tracts *en route* for the cerebral cortex are grouped together as the fillet tracts or lemnisci. These are the ***mesial lemniscus***, ***spinal lemniscus*** and ***trigeminal lemniscus***, all consisting of secondary neurones concerned with general body sensations, which end at the thalamus (on the side of the body opposite to that on which the primary neurones originate) ; and the *lateral lemniscus*, the afferent path for hearing, which consists of neurones terminating in the medial geniculate body. Third (or higher) order neurones complete the path from thalamus (or other intermediate station) to the cortex.

(1) **Tactile and Kinæsthetic Sensations.** (*a*) The Mesial Lemniscus represents the continuation of the *dorsal columns* of the cord, and therefore conveys conscious kinæsthetic and fine tactile impulses. These primary fibres, of rather large average size, pass up in the columns of Goll and Burdach, and, preserving their dermatomal lamination, end at the *nuclei Gracilis and Cuneatus* in the lower medulla, the ones from lower limbs at Gracilis, those from upper parts at Cuneatus (Fig. 196). The cells of these two nuclei give off secondary axons which cross the midline and pass upwards in the mesial lemniscus through pons and mid-brain and posterior limb of the internal capsule, still in a dermatomal order, to reach the postero-ventral nucleus of the thalamus. From this, tertiary fibres pass mainly to the sensory cerebral cortex (but some to the mesial and central nuclei of the thalamus).

(*b*) The spinal lemniscus is formed from the *two spinothalamic tracts.* The primary neurones of this path end in the grey matter of the cord, and its secondary axons, mainly small fibres, have already crossed in the cord. The fibres of the *lateral spinothalamic* tract, conveying pain and thermal sense, remain in a superficial position on the lateral aspect of the medulla, while those of the *ventral spinothalamic* tract, mediating light tactile sensations, lie more deeply. Ultimately both join the mesial fillet and terminate in the lateral nucleus of the thalamus, from which third order fibres proceed to the cerebral cortex (or mesial and ventral thalamic nuclei).

(*c*) The trigeminal lemniscus is derived from the sensory root of the Vth cranial nerve, the cell station being in the Gasserian ganglion. These primary neurones branch, forming ascending and descending " roots." These end in the sensory trigeminal nuclei, in pons, medulla and cervical cord, and from the cells of these arise the second order fibres which cross the midline at various levels and proceed with the two other lemnisci to the thalamus.

Afferent fibres of the VIIth, VIIIth (vestibular), IXth and Xth crania

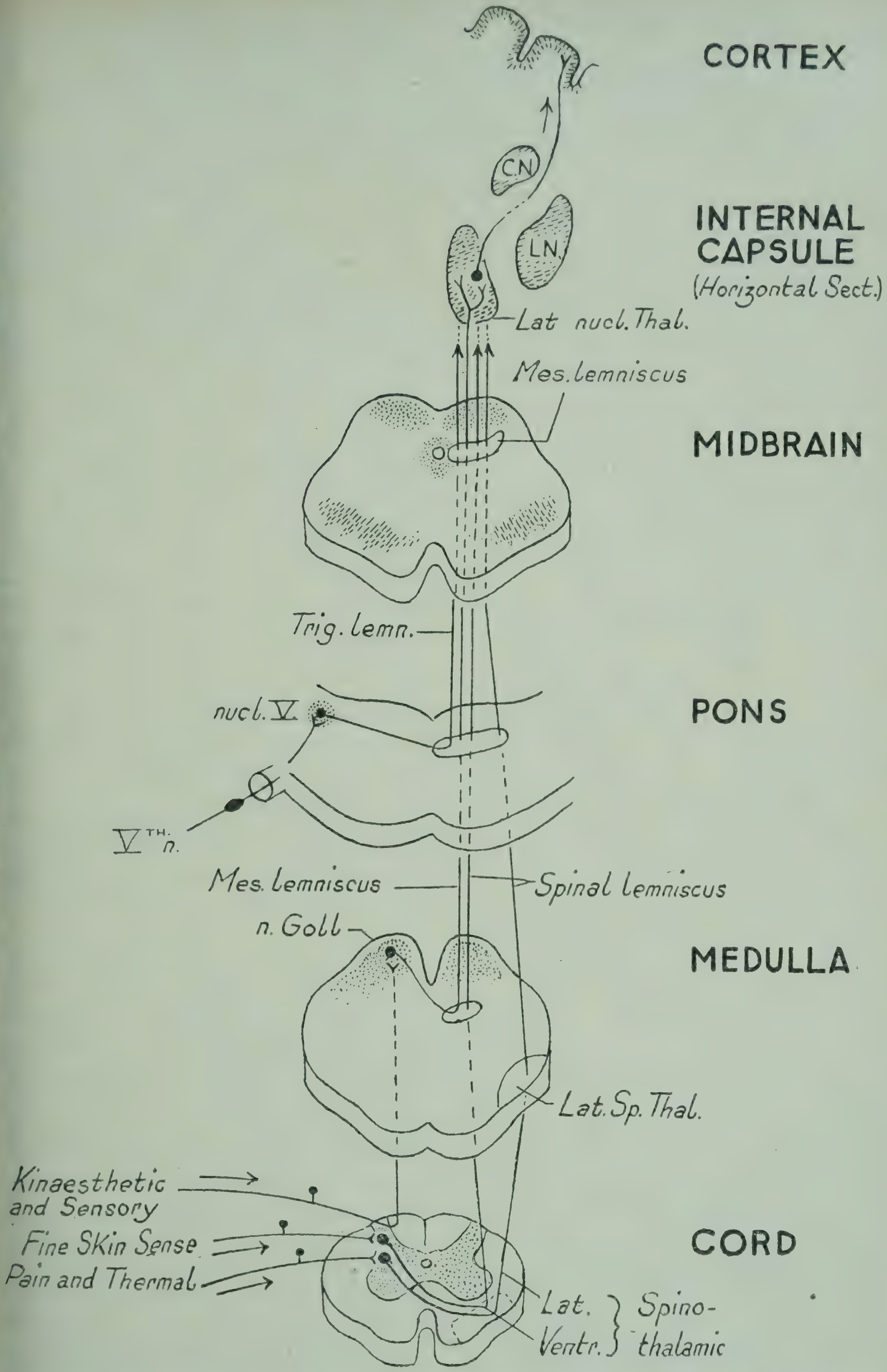

CORTEX
C.N.
INTERNAL CAPSULE
(Horizontal Sect.)
L.N.
Lat nucl. Thal.
Mes. lemniscus
MIDBRAIN
Trig. lemn.
nucl. V
PONS
V^TH n.
Mes. lemniscus
Spinal lemniscus
n. Goll
MEDULLA
Lat. Sp. Thal.
Kinaesthetic and Sensory
Fine Skin Sense
Pain and Thermal
CORD
Lat.
Ventr.
Spino-thalamic

FIG. 196. Diagram to illustrate main somatic sensory tracts. (n. Goll = *n. Gracilis* and *n. Cuneatus*.)

nerves also join the common fillet in a similar manner. Of the vagus fibres, for instance, about 70 per cent. are afferent, mostly unmyelinated, and with cell stations chiefly in the ganglion nodosum *; they have many important reflex functions.

---

* FOLEY and DU BOIS. *J. Comp. Neurol.*, 1937, **67**, 49.

(2) **Auditory Sensations.*** THE LATERAL LEMNISCUS is the path for auditory impulses. These arise in the cochlea and pass in by primary axons, with cells in the spiral ganglion, to form the cochlear division of the VIIIth cranial nerve. This divides round the restiform body, one branch ending in the dorsal, the other in the ventral cochlear nucleus.

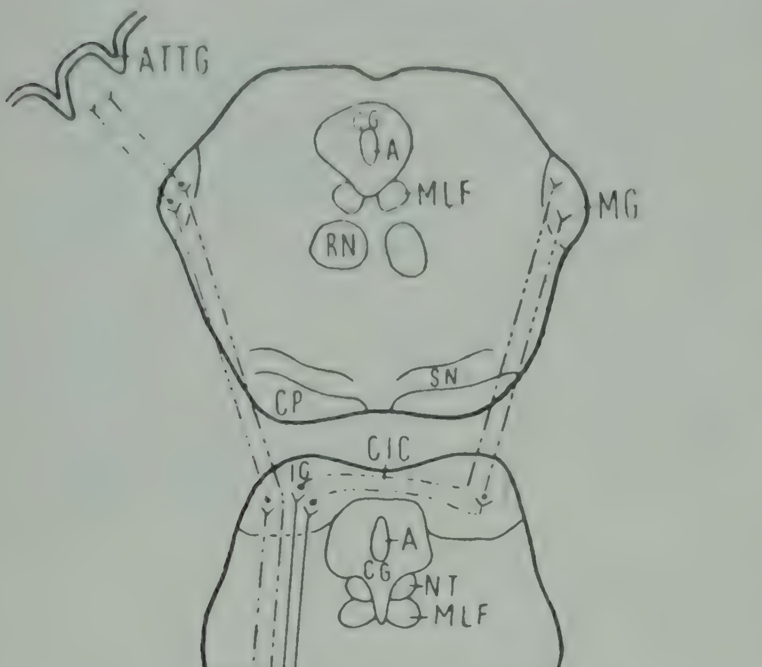


FIG. 197. Diagram of the central ascending auditory pathway in the monkey. VIIIc, cochlear nerve; CR, restiform body; DCN, VCN, dorsal and ventral cochlear nuclei; MLF, mesial longitudinal fasciculus; SO, superior olive; NLL, nucleus of lateral lemniscus; ML, mesial lemniscus; IC, inferior colliculus; CIC, commissure of inferior colliculus; MG, mesial geniculate body; A, aqueduct; CP, cerebral peduncle; ATTG, ant. tr. temporal gyrus. (After BARNES, MAGOUN and RANSON, *J. Comp. Neurol.*, 1943, **79**, 129.)

From these nuclei the fibres pass in three strands to a series of contralateral relay stations; these strands, dorsal, intermediate and ventral, cross the reticular formation, and reach the region of the superior olive (Fig. 197). On the way, some fibres of the dorsal strand end in the reticular formation. Having reached the superior olive some fibres end there in relays, and others pass up direct, with those relays, to the lateral lemniscus; many of these end in the nucleus of the lateral lemniscus, the remainder in the inferior colliculus; from here the next relay leads to the mesial geniculate body of the thalamus, and from this a final relay leads to the cortical auditory centre in the anterior transverse temporal gyrus. Some relays from the inferior colliculus cross back to the opposite inferior colliculus and so on by another relay to the other mesial geniculate body and cortex. This gives bilateral cortical representation and there are in addition some uncrossed paths from the superior olive upward which lead to the same result.

The cochlear path is thus effected by a chain of neurons with cells in cochlear nuclei, superior olive, lateral lemniscus nucleus, inferior colliculus, and mesial geniculate body, and each ear is connected to both sides.

There is a spatial projection of the different cochlear elements in the mesial geniculate body. It is claimed † that small localised lesions in the mesial geniculate body lead to loss of appreciation of particular pitches of sound, *e.g.* in the dorsal part to loss of high notes and in the ventral part to loss of low-pitched notes, *i.e.* particular fibres carry impulses aroused in discrete parts of the cochlea by particular notes. This point-to-point representation is continued to the cortex.

Auditory Reflexes. The inferior colliculus acts as a centre for auditory reflexes; from it arise the tecto-bulbar, tecto-spinal and other tracts, such as the posterior longitudinal bundle, by which auditory impressions are correlated with movements of head and eyes.

---

* For recent views *v.* BARNES, MAGOUN and RANSON. *J. Comp. Neurol.*, 1943, **79**, 129.

† ADES, METTLER and CULLER. *Amer. J. Physiol.*, 1939, **125**, 15.

Another important connection of the cochlear nerve is one *viâ* the superior olive to the oculomotor nuclei, and one direct to the VIth cranial nerve which is motor to the external rectus muscle. Hence on hearing a sound, the eye is reflexly turned to that side. In animals the incidence of sound on the ears causes erection of the pinnæ. This is effected by a reflex, the ingoing path of which consists of auditory nerve fibres, and the outcoming path of fibres of the VIIth nerve. Another auditory reflex is the contraction of the tensor tympani and stapedius muscles when sounds first fall on the ear; these muscles are innervated by a branch of the Vth cranial nerve. This reflex is bilateral even if sounds fall on one ear only. It is most marked for sounds of high pitch. Its object is to put the drum into correct tension for the perception of sounds.

(3) **Visual Sensations.** The rods and cones of the retina are visual sensory end-organs and neurones. The bipolar cells of the retina would then be the second order relay neurones, and convey the impulses to the third order neurones, or ganglion cells. The axons of these pass out from the eye as the optic nerve. Having entered the cranial cavity the nerve meets its fellow from the other eye, forming the optic chiasma, in which fibres partially decussate. The fibres thus form the optic tracts which travel round the crura cerebri and then divide to end in relation to fourth order neurones in three different places, viz. the pretectal region (a transitional area between thalamus and mid-brain), the superior colliculi, and the lateral geniculate body (Fig. 198). The first two are connected with reflex paths, the last with the transit of visual sensations to the occipital lobe.

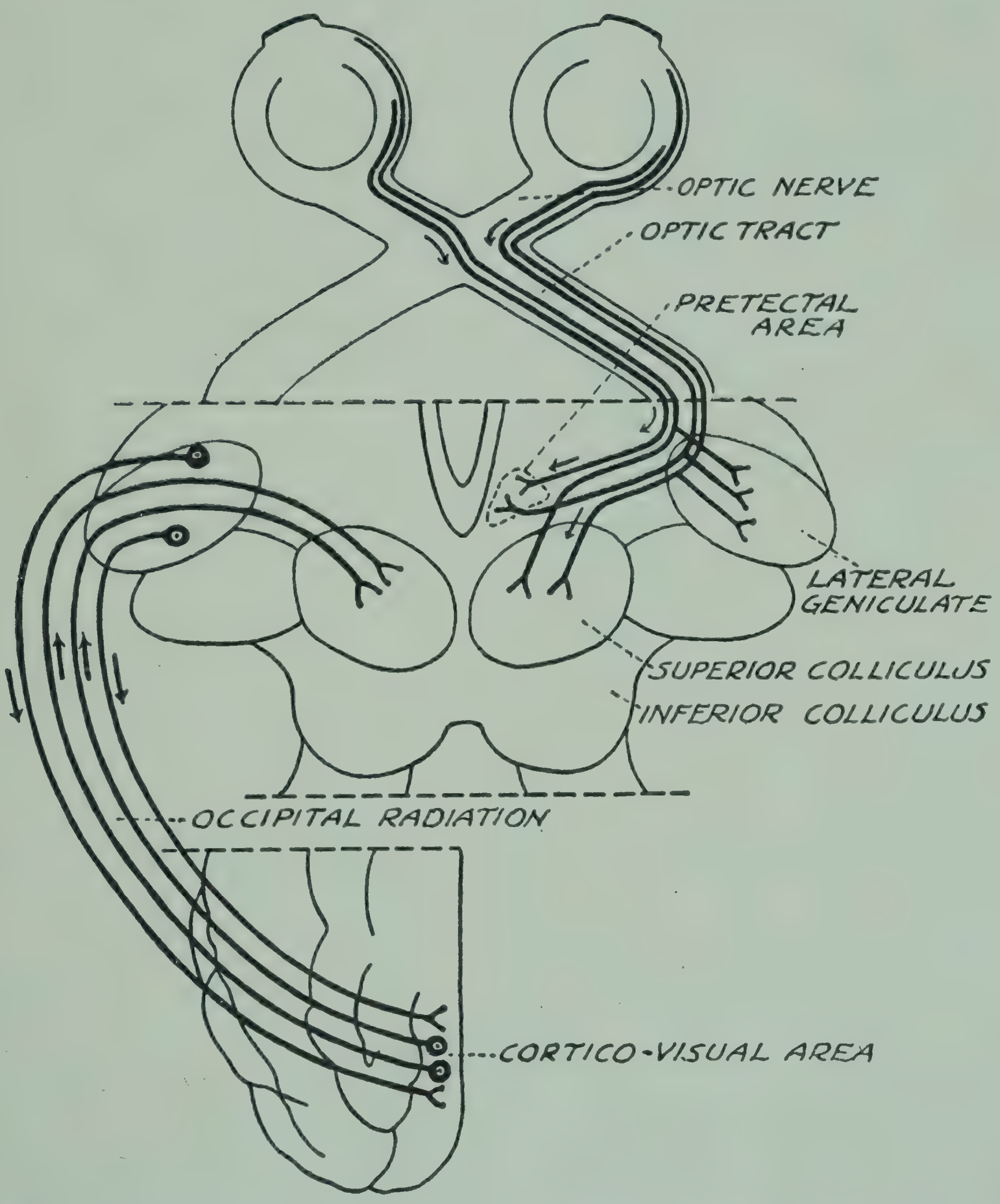


FIG. 198. Diagram of the Optic Paths. For clarity the first parts of the paths are shown in the right and the final stages on the left. (Modified slightly from BARRIS, *Arch. Ophthal.*, 1935, **14**, 61.)

The optic nerve may contain four different sets of fibres: (1) Those which convey visual impressions to the brain (coarse fibres); (2) those going to the reflex centres (fine fibres); (3) efferent fibres from the brain to the retinæ; (4) nerves travelling from one retina to the other.

The last two types, if they exist at all in mammalia, are of unknown function.

(1) Owing to partial decussation at the chiasma, the visual fibres of each optic tract are derived from both eyes, viz. from the temporal half of the retina and macula of the same side, and the nasal half of retina and macula of the opposite side; thus the right optic tract contains fibres derived from the right sides of both right and left retinæ and maculæ (Fig. 198). Each optic tract connects (for vision) with the dorsal nucleus of the lateral geniculate body (a specialised thalamic nucleus) on its own side, and, after relaying there, the impulses all pass onwards by the optic radiations to the visual centre in the occipital cerebral cortex.

By transneuronal degeneration it has been shown that there is clearly defined local representation of retinal areas in the lateral geniculate body.*

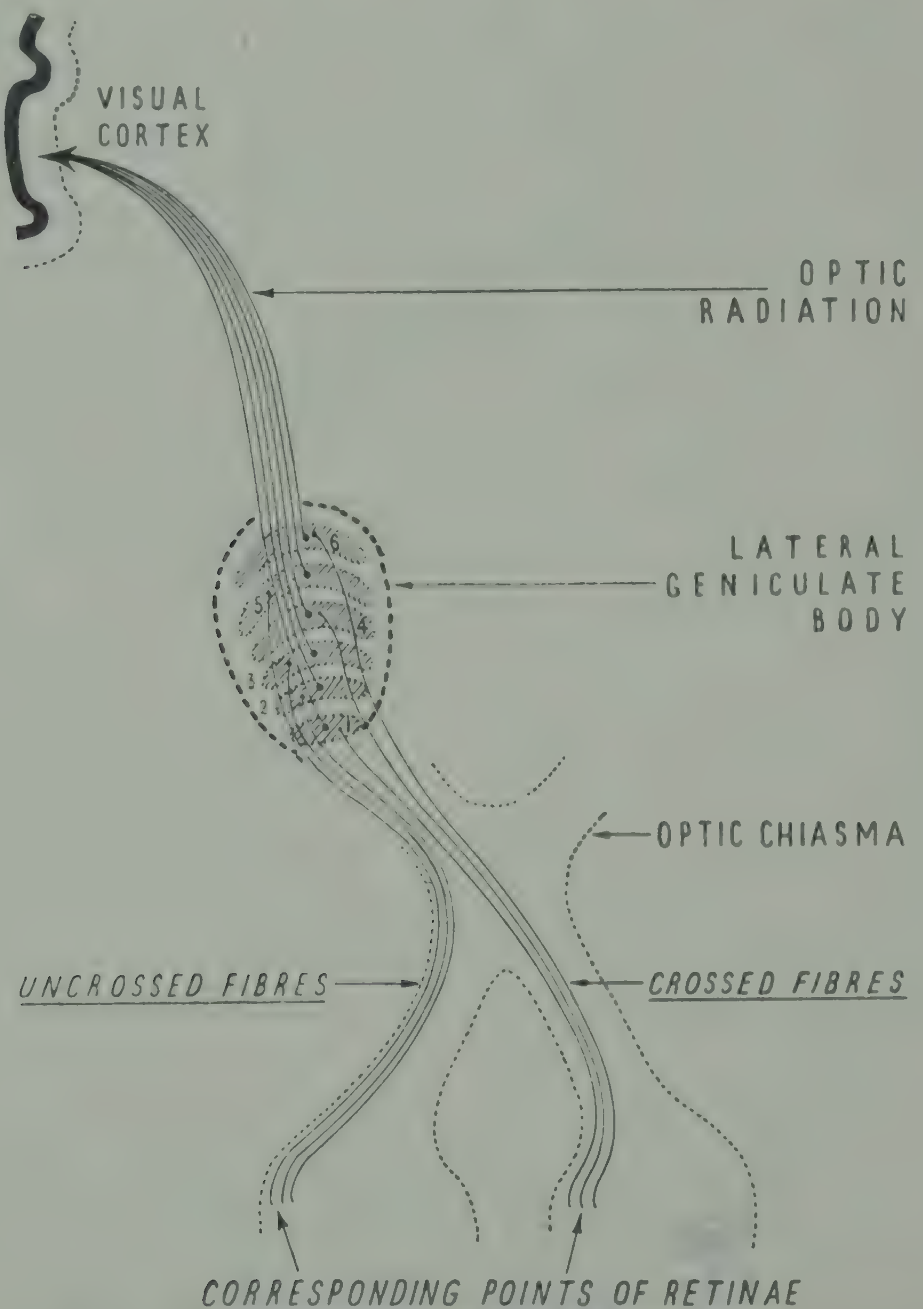


FIG. 199. To illustrate the convergence of impulses from corresponding retinal points at the visual cortex *viâ* the lateral geniculate body. (Modified from LE GROS CLARK.)

The macula has an especially large area of representation. There is a curious arrangement in the lateral geniculate body, which contains six cell laminæ. Fibres from corresponding points on the two retinæ end at vertically aligned cells in such a way that the uncrossed fibres from the temporal side of the retina end in layers 2, 3 and 5, while the crossed fibres from the corresponding points on the nasal side of the other retina end in layers 1, 4 and 6 (Fig. 199). The fourth order fibres from these cells representing corresponding points are then apparently brought to a focus at the visual cortex. It has been suggested that the 3-layer arrangement for each side may be related to colour vision.

---

* LE GROS CLARK and PENMAN. *Proc. R. S.*, 1934, **114B**, 291.

It will be seen that the occipital cortex is in this manner in connection with the halves of both retinæ of its own side, and impulses from corresponding points emerge close together. Since, as will be learnt later, each half of the retina receives light images of the opposite half of the whole visual field, this means that with vision, as with touch, each side of the brain receives impressions derived from the opposite side of the environing world.

(2) Some special afferent fibres of the optic tracts, and probably also some collaterals from the visual fibres, pass to the *superior colliculi* ; here they form the "stratum opticum" and terminate in the superficial layers : there is some degree of retinal area representation. This type of ending is concerned with reflex eye movements. It may represent a primitive vision centre.

Another reflex arc is that *viâ* the *pretectal region*, the pupillo-motor path described below.

It is of practical importance to be able to locate an injury to the visual nerve paths. Injury to an optic *nerve* causes blindness of the eye to which the nerve belongs, and exposure of the eye to light will not then elicit the pupil reflex. Injury to an optic *tract* causes blindness of the halves of *both retinæ on the same side* as the lesion, that is to say, *blindness to external objects on the opposite side* to the injury. Injury at the chiasma, *e.g.* by pituitary tumours, may only interfere with the fibres from the nasal halves of both retinæ, so causing loss of the temporal fields of vision—*bitemporal hemianopia.*

It is interesting to note that, whereas, in most other nerve paths, crossing of the fibres occurs from one side to the other as they travel to the brain, so that the left side of the brain is connected only with the right side of the body, this is not the case with the optic—or auditory—impulses. Injury to the optic radiation of the calcarine cortex will cause blindness of the corresponding halves of *both* retinæ, *i.e.* the half of each retina on the same side (and so blindness to the external world on the *opposite* side), but will not affect the pupil reflex, because the fibres concerned have already turned aside to go to the pretectal region.

EYE REFLEXES. *The Correlation of Head and Eye Movements.* These may be of two kinds, compensatory and associated. The compensatory ones take place when the head is rotated in one direction and the eyes in the opposite direction, in order still to fix the gaze on some stationary object.

The associated movements occur when the head and eyes rotate in the same direction in order to fix the gaze on an object which is moving round the observer. Both these associated movements appear to be effected *viâ* the superior colliculi. These receive not only afferent fibres from the optic tracts, but also efferent ones from the occipital cortex (Fig. 198). From the superior colliculi arise the motor tecto-bulbar and tecto-spinal tracts. Probably also co-ordination is effected through the inferior olivary bodies which connect with the oculomotor nuclei *viâ* the central tegmental tract, with the upper cervical cord by the olivospinal tract, and with the cerebellum.

The Correlation of Limb and Eye Movements. For the protection of the eyes there is an important reflex which can be initiated principally by vision, but also by the other distant receptors, such as hearing. This involves (*a*) the firm closing of the lids, and, in addition, where necessary, the raising of the arm in front of the head (or the hands in front of the eyes) ; (*b*) the flexing of the neck and spine so that the face approaches the abdomen ; (*c*) lastly, the reflex may cause the raising of the legs flexed at the knees so that they form an additional protection for the face. This reflex operates through the tecto-spinal tract.

The Pupillary Reflex. The plain muscle of the pupil is supplied by two nerves, viz. the sympathetic which causes dilatation, and the IIIrd cranial, which causes constriction. When light falls on the eye, the pupil constricts —the light reflex. This is effected by a reflex from the pretectal region, in which special fine fibres from the optic nerve terminate : from here relay fibres pass to end at the IIIrd nerve nucleus, from which axons pass out to the constrictor pupillæ muscle. There is some crossing over of neurones from side to side in the pretectal region, which explains the fact that when one eye is exposed to light the pupils of both eyes contract (Fig. 200).

In some disorders the phenomenon of the *Argyll-Robertson pupil* is encountered. In this, the pupil does not react to light, yet the IIIrd nerve nucleus must be intact, because the pupil still contracts during accommodation of the eye for near vision. The lesion is probably in the pretectal region ; destruction of the superior colliculi, which was formerly believed to be the relay station, does not affect the light reflex,* but destruction of pretectal region as well as of the superior colliculi does.†

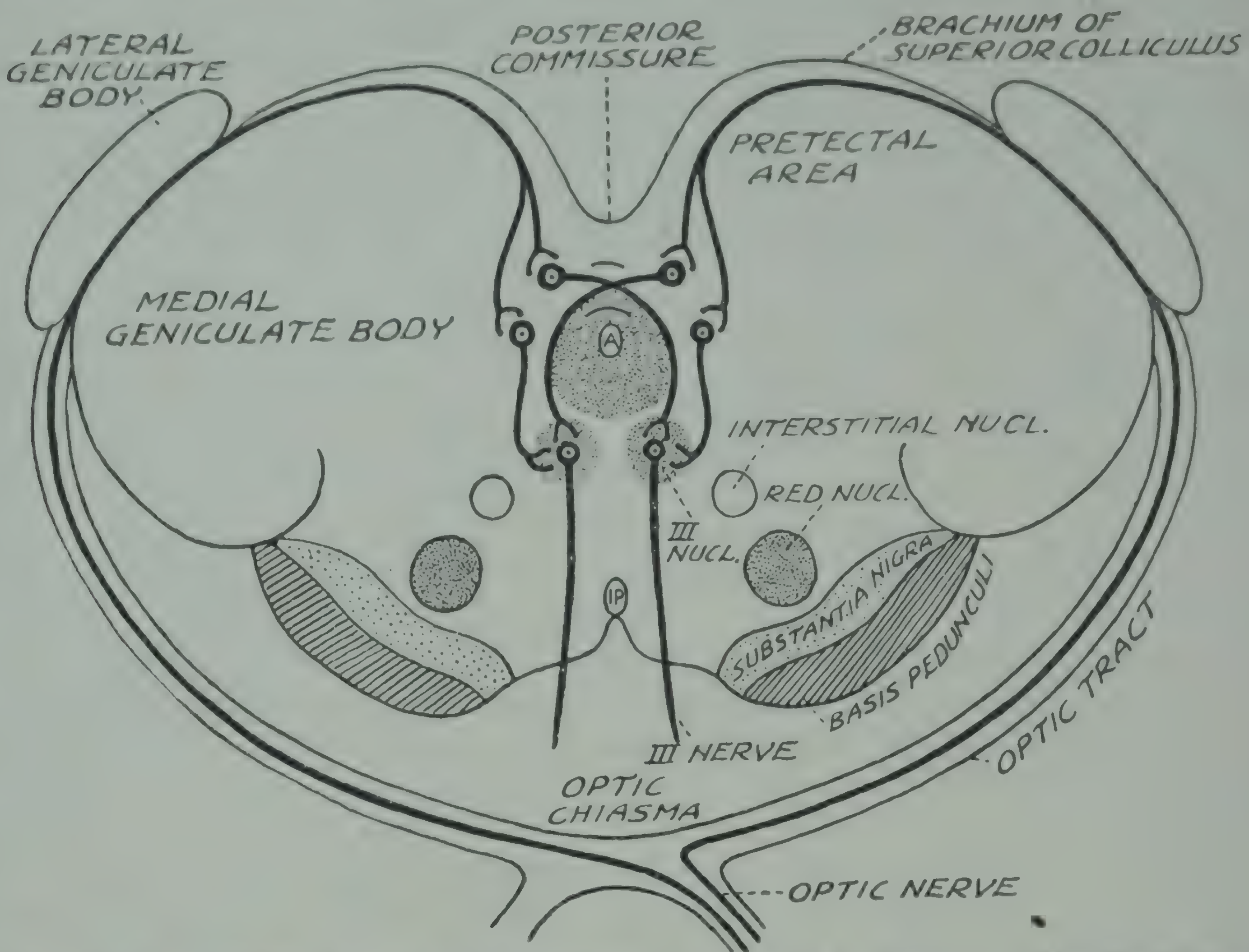

LATERAL GENICULATE BODY
POSTERIOR COMMISSURE
BRACHIUM OF SUPERIOR COLLICULUS
PRETECTAL AREA
MEDIAL GENICULATE BODY
A
INTERSTITIAL NUCL.
RED NUCL.
III NUCL.
IP
SUBSTANTIA NIGRA
BASIS PEDUNCULI
OPTIC TRACT
III NERVE
OPTIC CHIASMA
OPTIC NERVE

Fig. 200. Diagram of the Path Concerned in the Light Reflex. (Modified from Magoun and Ranson, *Arch. Ophthal.*, 1935, **13,** 791.)

(4) The Paths for Taste and Smell. The *nerves for taste* are the VIIth and IXth cranial, the cells of origin being in the geniculate and petrous ganglia respectively. These primary neurones end in the sensory nuclei of the VIIth and IXth nerves, from which second order fibres run, probably with the common fillet tracts of the opposite side, to the thalamus. From this, third order neurones continue to the cortical centre.

The *nerves for smell* perforate the cribriform plate and enter the olfactory bulb, of complex structure, where a relay occurs. The relaying fibres form the olfactory tracts. Ultimately the impulses reach the uncinate gyrus, the hippocampus and other structures which together form the rhinencephalon or olfactory brain.

THE THALAMUS

With the sole exception of the olfactory path, all the afferent paths which are travelling to the cerebral cortex converge on to the thalamus, which is their chief relay station. Thus, it receives optic and auditory radiations, the spinal, mesial, and trigeminal lemnisci and cerebellar fibres. It also has connections with various subcortical stations. All cortical connections are with the cortex of the same side.

The thalamus contains many nuclei, which are differently described by different authorities. In mammalian forms, it may be divided, according to Le Gros Clark ‡ into an upper and a lower level. *The upper level* is dorsally placed, and contains, among others, the lateral and dorso-medial nuclei, the supra-geniculate and centre median nuclei. These upper levels do *not* receive direct afferent fibres, and are related to the association areas of the cortex ; they also have connections with other nuclei of the diencephalon. It is probable that the more ventrally placed ones of this group also have functions in connection with visceral control, and so connect with the hypothalamus.

---

* Keller and Stewart. *Amer. J. Physiol.*, 1932, **102,** 64.

† Magoun and Ranson. *Arch. Ophth. N.Y.*, 1935, **13,** 791, 862 ; Barris. *Arch. Ophth. N.Y.*, 1935, **14,** 61.

‡ Le Gros Clark. *Brain*, 1932, **55,** 406 ; Le Gros Clark and Boggon. *Phil. Trans.*, 1935, **224B,** 313.

The lower levels of the thalamus are chiefly ventral, and receive afferent impulses from the brain stem, spinal cord and hypothalamus, and relay these to the upper level nuclei and to the cerebral cortex. They include (1) the main ventral nucleus and the medial ventral nucleus, (2) an anterior group, (3) an optic tract group, including the dorsal and ventral nuclei of the lateral geniculate body, the large-celled nucleus of the optic tract, and the pretectal nucleus, (4) the medial geniculate body.

Physiologically * there are three main groups of thalamic nuclei :

(i) Nuclei *with purely subcortical connections* (chiefly the anterior ventral groups), which do not degenerate after hemidecortication. They are the seat of primitive sensations, and probably serve to connect together the various nuclei of the diencephalon and have certain functions in visceral control.

(ii) Nuclei which are *relay stations on the afferent path* to the cortex. These are chiefly ventral, and include the geniculate bodies and the posteroventral and lateroventral nuclei.

(iii) *Association Nuclei.* These are the dorso-medial and lateral nuclei and pulvinar. They connect with other nuclei of the diencephalon (thalamus, hypothalamus and subthalamus) and also with the parietal and anterior frontal association areas of the cerebral cortex, but receive no afferent fibres from the sensory systems.

The chief function of the thalmus is undoubtedly that of receiving, modifying and relaying sensory impressions to the cortex. There is also important reciprocity between cortex and thalamus, a nucleus sending fibres to the cortex likewise receiving fibres from it. It has been shown that after receipt of sensory impulses at the cortex, an after-discharge occurs, which can be detected by a suitable amplifier, and traced back to the thalamus, but not further back.† The thalamus and cortex thus work in close collaboration, some parts of the thalamus being connected with extensive areas of cortex, and there is little doubt that the organisation of the cerebral cortex is largely effected by the intermediation of the thalamus.

Removal of parts of the cerebral cortex leads to retrograde degeneration of the thalamic relay nuclei connected with those parts. Thus it can be shown that the lateroventral parts project to areas 4 and 6 (*v.* p. 299) of the cortex and are concerned with unconscious kinæsthetic afferents, while the posteroventral parts act as relay stations for conscious sensations arriving *viâ* the lemnisci on their way to the cortex. Hence destruction of the posteroventral portion of the thalamus causes disturbances of skin and kinæsthetic sensations on the opposite side. The representation of areas is demonstrable in the thalamus, face being in the most mesial, arm and leg respectively more lateral parts of the posteroventral nucleus. There is, however, no such segregation of modalities of sensation as is seen in the cord, since the mesial lemniscus and spinal lemnisci terminate at about the same parts.

The lateral geniculates, as we have already seen, are relay stations on the visual path, and the mesial geniculates on the auditory and vestibular paths. The pulvinar serves to correlate visual and auditory impressions.

Disease of the thalamus is often accompanied by extreme hypersensitiveness in the areas represented. This is probably a release phenomenon due to severing of cortical connections. There seems little doubt that the thalamus is capable of sensation of a crude type independently of the cortex. This explains why some responses, especially to pain and temperature, may remain even after the removal of large areas of the cerebral cortex.

We have already seen that the decorticate animal exhibits a tendency to a display of " sham rage " on receipt of almost any stimulus, and this is believed to be due to the removal of the normal control of the cortex, by which such emotional expressions are toned down. It is probably in the upper levels of the thalamus that afferent impressions become loaded with an emotional quality.

II. UNCONSCIOUS AFFERENTS

The paths used for conscious impressions may also often be traversed by impulses that do not emerge into consciousness, but are concerned wholly with reflex effects. Some paths, however, are devoted wholly to the transmission of impulses which normally arouse no sensations. If, by exceptionally strong stimulation, such impressions do force their way into the field of consciousness, the sensations aroused are vague in character and lacking in local sign. The most important of these tracts are concerned with unconscious kinæsthetic impressions, related to the position and movements of the limbs, neck or trunk. Their most important destination is the cerebellum. Others lead from the viscera, and have connections with the spinal and bulbar centres. Among these paths the following are important :—

---

* Earl Walker. *J. Nerv. Ment. Dis.*, 1937, **85**, 249.

† Adrian. *J. Physiol.*, 1941, **100**, 159.

(1) **The spinocerebellar tracts** arise from the secondary neurones of the Clarke-Stilling column in the spinal cord. The large axons of these cells pass into the *dorsal and ventral spinocerebellar tracts*, chiefly of the same side and ascend through the medulla. Here the dorsal spinocerebellar tract takes a more dorsal position and forms an important part of the restiform body, or inferior cerebellar peduncle, with which it travels to, and ends in, the cortex of the vermis of the cerebellum.

Other important afferents originate from the dorsal columns of the cord, and reach the cerebellum from the bulbar nuclei, chiefly of the opposite side, *viâ* the restiform body. Among these are axons (the arcuate fibres) from the nuclei Gracilis and Cuneatus, which relay unconscious kinæsthetic impulses ascending the spinal cord.

(2) **Vestibular Paths.** As would be expected, the vestibular nerve has extensive central connections concerned with the co-ordination of motor responses. It has synapses with cells in one of the four subdivisions of the vestibular nucleus, viz. medial (Schwalbe), lateral (Deiters), superior (Bechterew), or the vestibulo-spinal nucleus. From the cells of these nuclei four main tracts arise :—

(i) *Connects with the posterior longitudinal* bundle of both sides.
(ii) The *vestibulo-spinal tract*, which descends homolaterally in the cord.
(iii) *The vestibulo-reticular tract*, which connects with motor nuclei and reticular formation, of opposite side.
(iv) *The vestibulo-cerebellar tract*, which includes some direct fibres from the vestibular nerve. It is homolateral.

In addition to these tracts leaving the vestibular nucleus, there is the cerebello-vestibular or uncinate fasciculus, which descends to the vestibular nucleus from the roof nuclei of the cerebellum of the same side.

(3) **Olivo-cerebellar Tracts.** The olivary bodies have afferent connections with the spinal cord, thalamus, etc., and send out fibres which pass up in the inferior peduncles to the cerebellum.

(4) **Visceral Afferents.** These will be considered later.

For Reference

Bishop (1946). Neural Mechanisms of Cutaneous Sense. *Physiol. Rev.*, **26**, 77.
Hartline (1941). The Neural Mechanisms of Vision. (*Harvey Lects.*, 1941–42 ; p. 39.
Le Gros Clark (1942). The Visual Centres of the Brain and their Connexions. *Physiol. Rev.*, **22**, 205.
McCulloch (1944). The Functional Organisation of the Cerebral Cortex. *Physiol. Rev.*, **24**, 390.
Walker (1938). The Primate Thalamus. (Univ. Chicago Press.)

CHAPTER XX

THE CEREBRAL HEMISPHERES AND THE EFFERENT PATHS

THE cerebral hemispheres form the largest and most important part of the human brain. It is to the progressive evolution of this part that is due the rise in type in vertebrates. In development they are formed as two diverticula from the front part of an outgrowth of the first cerebral vesicle. In the lowest vertebrates these outgrowths are connected entirely with the olfactory sense organs, and we may regard the olfactory brain as a fundamental structure on which has been built up all the rest of the cerebral hemispheres. With the increasing importance of the visual sensations in the bony fishes there is still very little corresponding growth of the fore-brain, most of the fibres from the optic nerves going to the roof of the mid-brain (the optic lobes). The emergence of the cerebral hemispheres is associated with the development of nervous tissue in the roof of the fore-brain. At its first appearance this higher brain material still receives chiefly olfactory impressions. But the structure of the cerebral cortex thus laid down differs from that of the centres forming the brain stem or the olfactory lobe itself in that it provides for a very rich association of impulses between all its parts. The fibres entering the cortex break up into a fine meshwork of fibres which run tangentially to the surface and come in relation with innumerable dendrites of nerve cells situated at some little distance below the surface. We have here the first germ of an apparatus in which the nerve paths can be determined by education. In the amphibian and reptile, the cerebral cortex extends over the whole roof of the cerebral hemispheres, though even here a very large proportion of it is devoted to the association of olfactory impulses. The wider reactive powers of birds are based chiefly on an enormous development of the corpus striatum, whereas in mammals the chief development occurs in the roof of the cerebral hemispheres. With the increased entry of fibres, carrying impulses from all the other sense organs of the body, the olfactory part of the brain diminishes in importance, and in the higher mammals and man is altogether overshadowed by the newly formed structures of the cerebrum. The cortex of the olfactory brain is sometimes called the *allocortex*, and that of the remainder of the cerebrum the *isocortex*.

The whole surface of the brain is thrown by fissures into convolutions, by which means about a threefold increase of the surface grey matter is obtained. By the larger fissures the brain surface is divided into lobes. The chief lobes are (*a*) the frontal; (*b*) the parietal; (*c*) the occipital; (*d*) the temporal; (*e*) the limbic area. The chief fissures separating these are (*a*) the central or fissure of Rolando, between frontal and parietal lobes; (*b*) the Sylvian fissure between the parietal and temporal; (*c*) the parieto-occipital; (*d*) the calloso-marginal, forming the outer boundary of the limbic area. Each of the main lobes or areas mentioned above is further subdivided by numerous minor fissures.

MINUTE STRUCTURE OF THE CEREBRAL CORTEX. The cortex of the cerebral hemispheres consists of a layer of grey matter covering a deeper mass of white fibres. The white fibres are axons passing out of or into the grey matter.

Even the naked eye appearance of the grey matter, in vertical section, differs from place to place. Histologically the grey matter consists of layers of nerve cells embedded in neuroglia and nerve fibres, both medullated and non-medullated. The nerve cells vary in size and shape; one cell typical of the cerebral cortex is the *pyramidal cell*, a cone-shaped cell with one large apical process which runs towards the surface, various dendrites which are given off from the sides and lower angles of the cell, and another process (usually the axon) from the base of the cell which passes downwards giving off small collaterals in its course. Some of their axons (*projection fibres*), including those which enter the pyramidal tracts, pass by the corona radiata through the internal capsule and into the crura cerebri. Others, or their collaterals, may pass into the adjacent regions of the cortex, or by the corpus callosum into the opposite cerebral hemisphere (*association fibres*). Other neurones belong to Golgi's type II and are confined to the grey matter; they are called "*intracortical neurones*."

The structure of the allocortex (olfactory) * is totally different from that of the isocortex and cannot be considered here.

---

* LORENTE DE NÓ. *J. Physiol. Neurol.*, 1934, **45**, 381; **46**, 113.

In the isocortex, histological examination of the layers of cells and fibres shows that there are many different variations of a presumably common plan of stratification, and much labour has been devoted to an attempt to unfold this basic plan. When the whole cortex of man is systematically studied it is found that it can be divided into numerous areas, each of which has a distinctive overall structure, as judged by the thickness of the various layers of fibres and cells and by special modifications which they show from one area to another. The number of fields into which the whole cortex is thus divisible varies according to the criteria adopted by different investigators, and has been increased as a result of detailed study from an initial twenty areas * to fifty,† or even, according to some, to more than 200, all of which are presumed to have different functions.

Some of the more generally accepted structurally distinguishable areas, according to Brodmann, are shown in Fig. 201, and it is common to refer to the areas by the numbers there given.

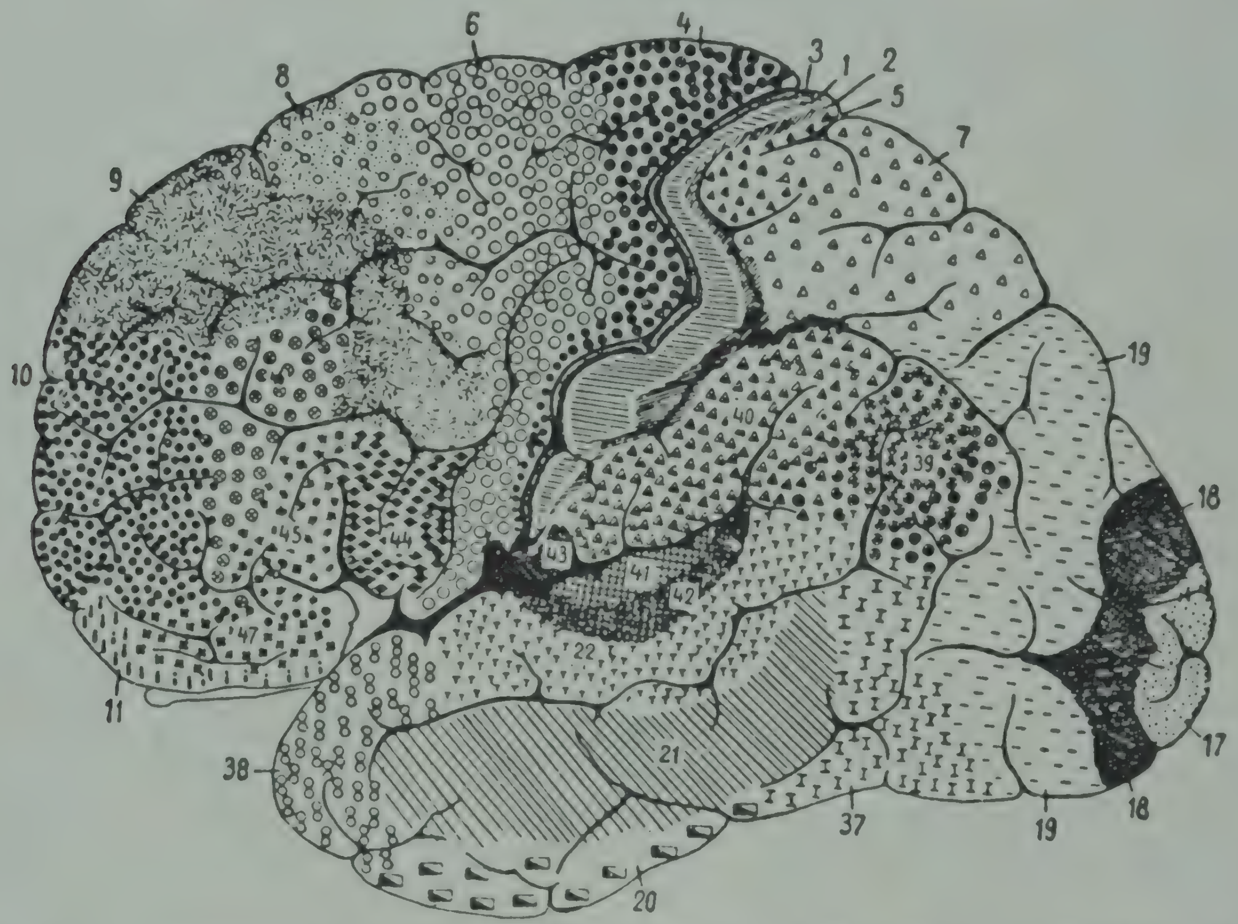


FIG. 201. Brodmann's Cytoarchitectonic Map of the Human Cortex. (From FULTON'S "Physiology of the Nervous System," Oxford Univ. Press.)

The plan of structure of the important parts, represented by the parietal, temporal and occipital lobes, generally agreed upon is that of a fundamentally six-layered stratification (I–VI) of cells of different types, layers IV., V. and VI., however, being often subdivided. The cell layers described are recognised by staining the cells for Nissl's granules, *e.g.* with methylene blue. If the medullated fibres are stained, as by the Weigert-Pal method, instead of the cells, new strata (1–6), consisting of fibres, are revealed (Fig. 202).

These six cell layers, in two laminæ, are :—

EXTERNAL LAMINA

I. The plexiform layer. It contains few cells, and is composed principally of the dendritic endings of cells belonging to deeper layers. It contains a few horizontal cells of Cajal, with processes which run approximately parallel to the surface.

---

* CAMPBELL. "Histological Studies on the Localisation of Cerebral Function," Cambridge University Press, 1905.

† ECONOMO. "The Cytoarchitectonics of the Human Cerebral Cortex." London, Oxford Univ. Press, 1929 ; BRODMANN, *Vergleichende Lokalisationslehre der Grosshirnrinde*, etc. Leipzig, Barth., 1909 and 1925.

II. The layer of small pyramidal and small granule cells.
III. The outer layer of medium pyramidal cells.
IV. (*a*) The layer of star pyramids. (The outer line of Baillarger (fibres) is at this level.)
IV. (*b*) The layer of stellate cells. Sometimes this layer is called the middle stellate or granule layer.

INTERNAL LAMINA

V. The inner layer of large deep pyramids. It is in this layer of the motor area that the giant Betz cells are found. (The fibre component of this layer, 5*b*, is the inner line of Baillarger.)
VI. The fusiform cell layer.

Internal to layer VI is the white matter of the cerebral cortex.

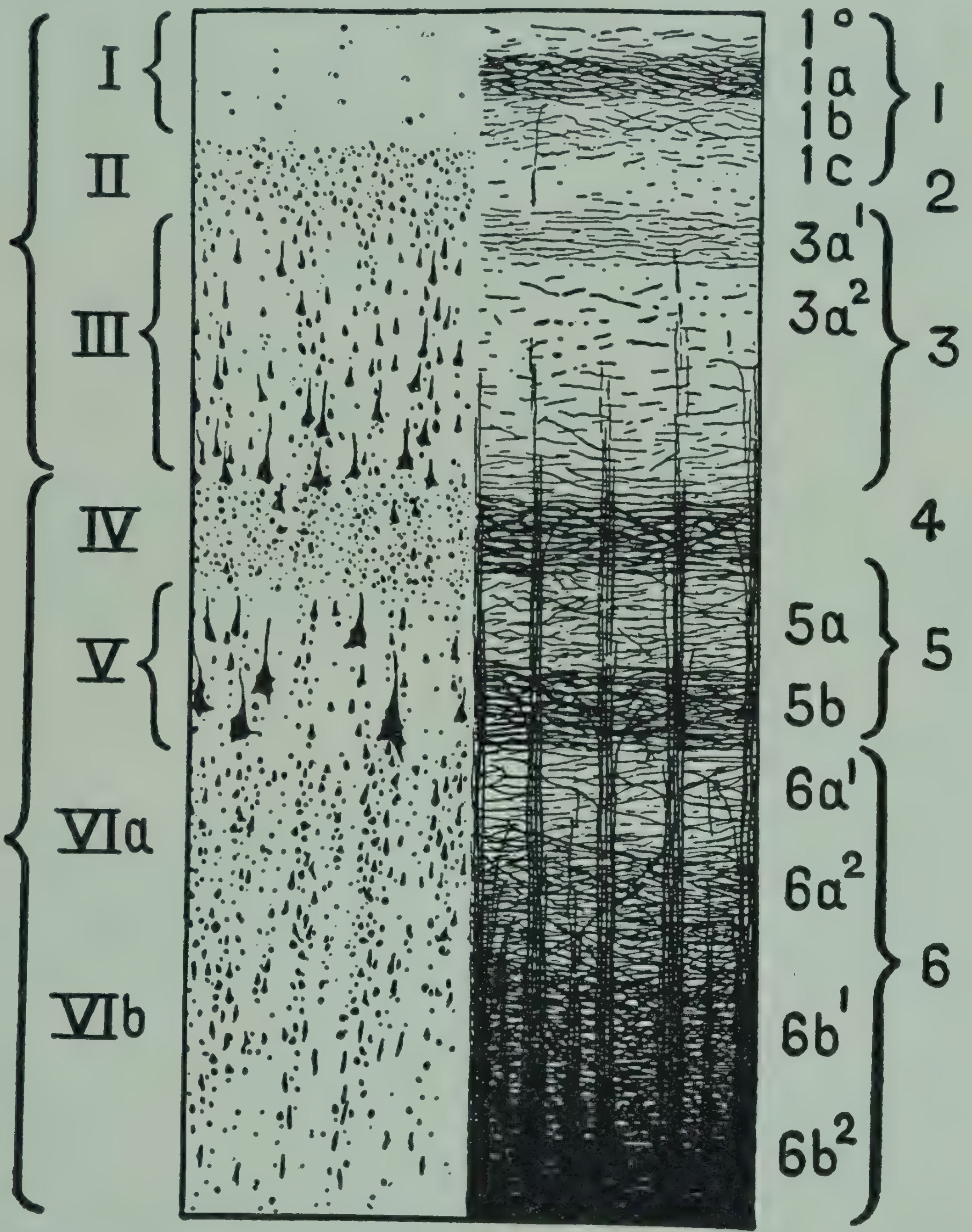


FIG. 202. Diagrammatic Representation of the Layers of the Human Cerebral Cortex according to Brodmann and Vogt. On the left the six cell layers. On the right the fibre layers, Layers 4 and 5b being respectively the outer and inner lines of Baillarger. (After ECONOMO, "Cytoarchitectonics of the Human Cerebral Cortex," Oxford Univ. Press, 1929.)

The pyramidal cells of all sizes possess, in general, three different kinds of fibre : an apical one, which runs towards the external surface, and then branches tangentially to the surface ; a basal one (axon) running towards or into the white matter, where it becomes medullated and runs a variable distance according to its function ; and a series of transverse collaterals, which run tangentially. In consequence, these tangential fibres form well-marked lines, particularly at layers IV and V. In consequence, if sections are stained by a method like Weigert's, sheaves of radial fibres will be seen running from the white matter towards the surface, giving off a rich meshwork of fibres to the intervening portions of the grey matter, while the bands of tangential fibres

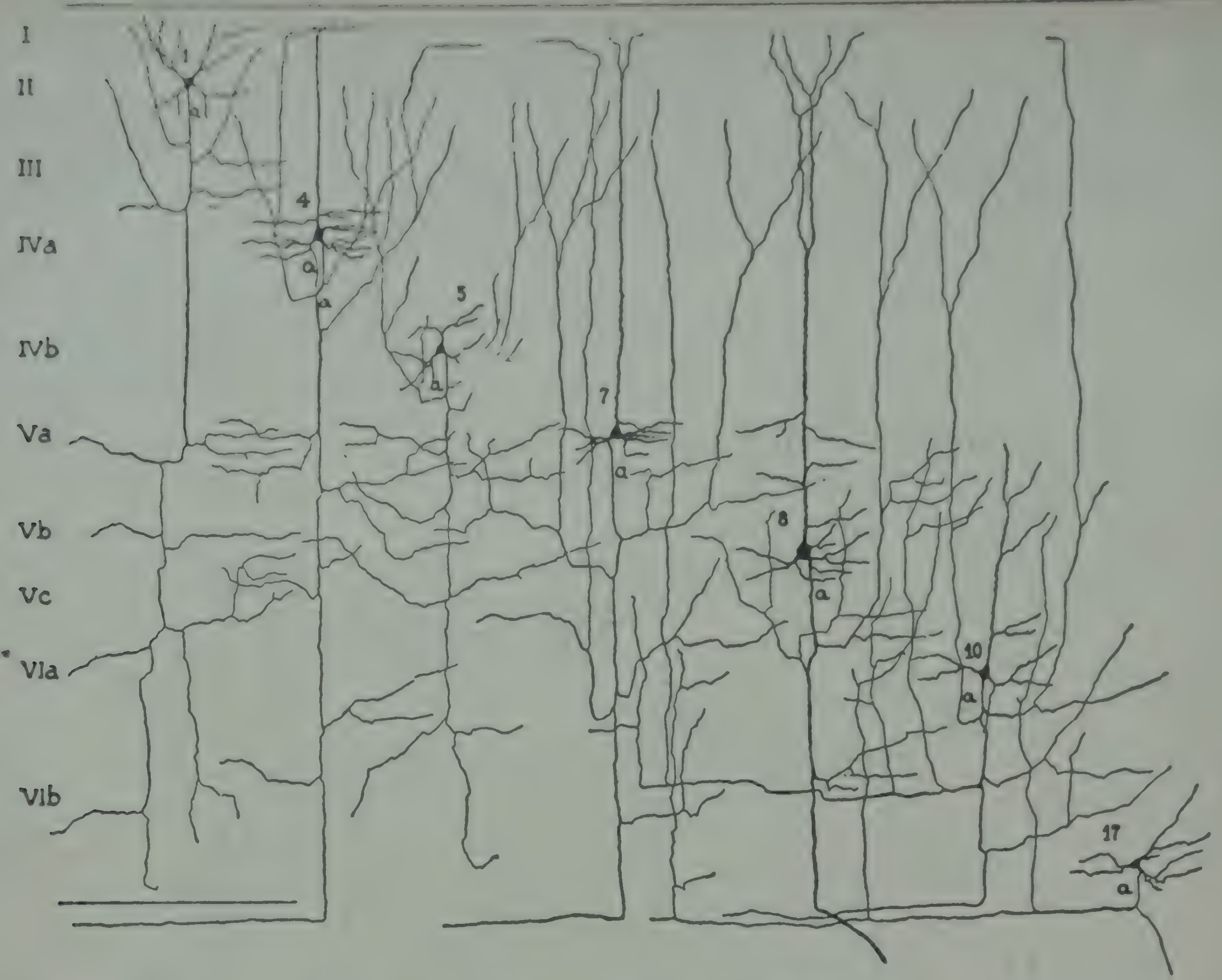

I
II
III
IVa
IVb
Va
Vb
Vc
VIa
VIb
1
a
4
a
a
5
a
7
a
8
a
10
a
17
a

FIG. 203. Form and Position of Descending-axon Neurones of Cerebral Cortex. Axons marked *a*. Note that collateral branches are concentrated in layers I–III and V–VI. Axons of cells 1 and 5 are confined to the cortex; those of 4, 7, 10 and 17 are association fibres; that of 8 is a projection fibre. (LORENTE DE NÓ, in FULTON's "Physiology of the Nervous System," Oxford University Press.)

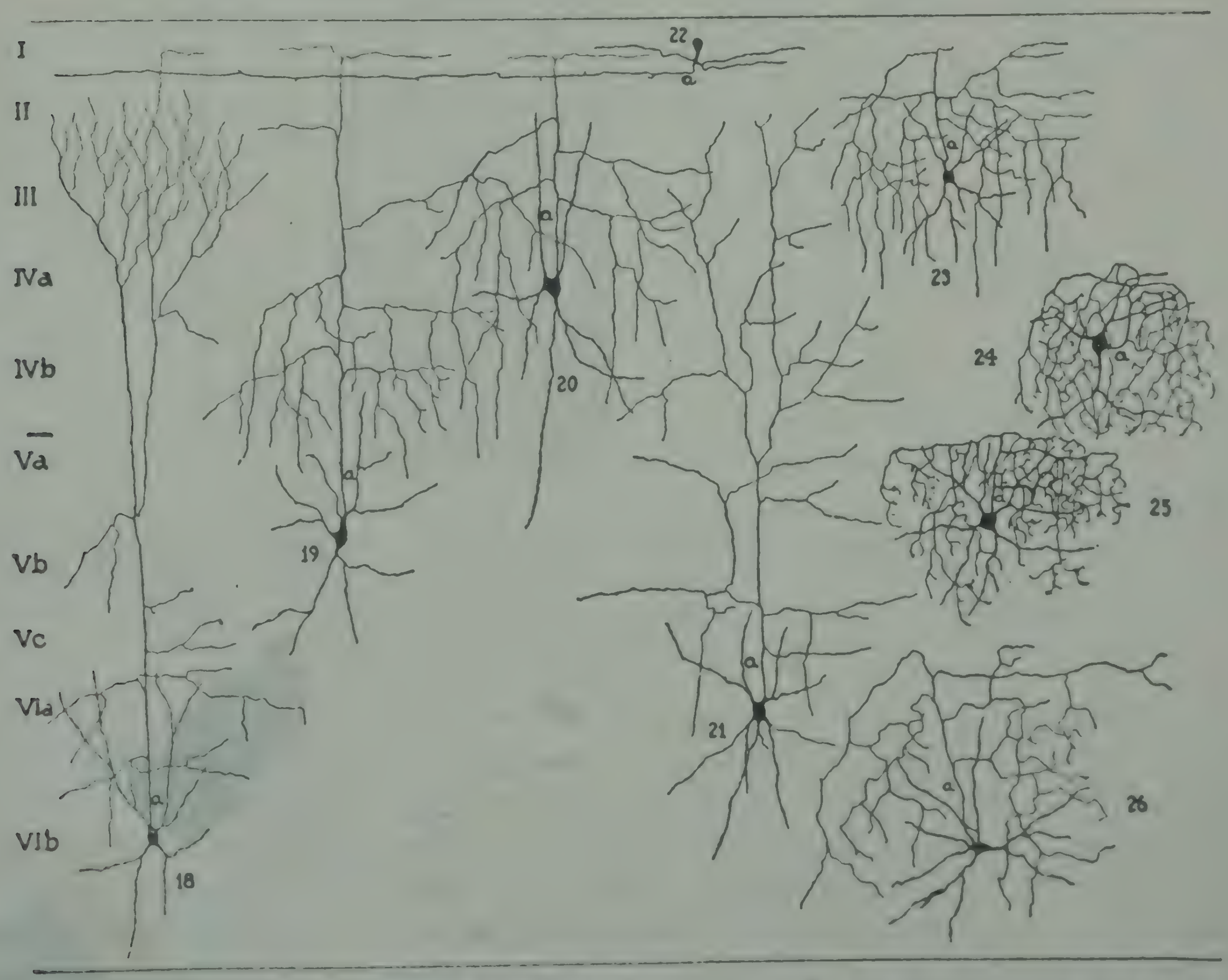

I
II
III
IVa
IVb
Va
Vb
Vc
VIa
VIb
22
a
a
20
23
a
24
a
19
a
25
a
21
a
26
a
18

FIG. 204. Types of Cerebral Cells with Intracortical Axons: axons marked *a*: 18 to 21, cells with ascending axons; 22, cell with horizontal axon; 23 to 26, cells with short axons. (LORENTE DE NÓ, in FULTON's "Physiology of the Nervous System," Oxford University Press.)

are seen running parallel to the surface. Thus we find (Fig. 202): (a.) The tangential fibre layer which corresponds to the plexiform layer; (b.) a layer of tangential fibres (4) at layer IV, corresponding to the stellate cell layer. This is sometimes called the outer line of Baillarger; (c.) an inner line of tangential fibres (5b) at layer V, corresponding to the layer of large deep pyramidal cells, sometimes called the inner line of Baillarger.

Two of the most clearly marked differences in areas are: (1) In area 4, the "agranular isocortex," a motor area, layer V is occupied by giant pyramidal cells, the axons of which enter the pyramidal tract. The largest of them are called the giant Betz cells. (2) In area 17, the visual area, the pyramidal cell layers are thin, the large pyramidal cells being very few. The lines of Baillarger are very distinct, particularly the outer, which is sometimes called the line of Gennari, a characteristic of this region.

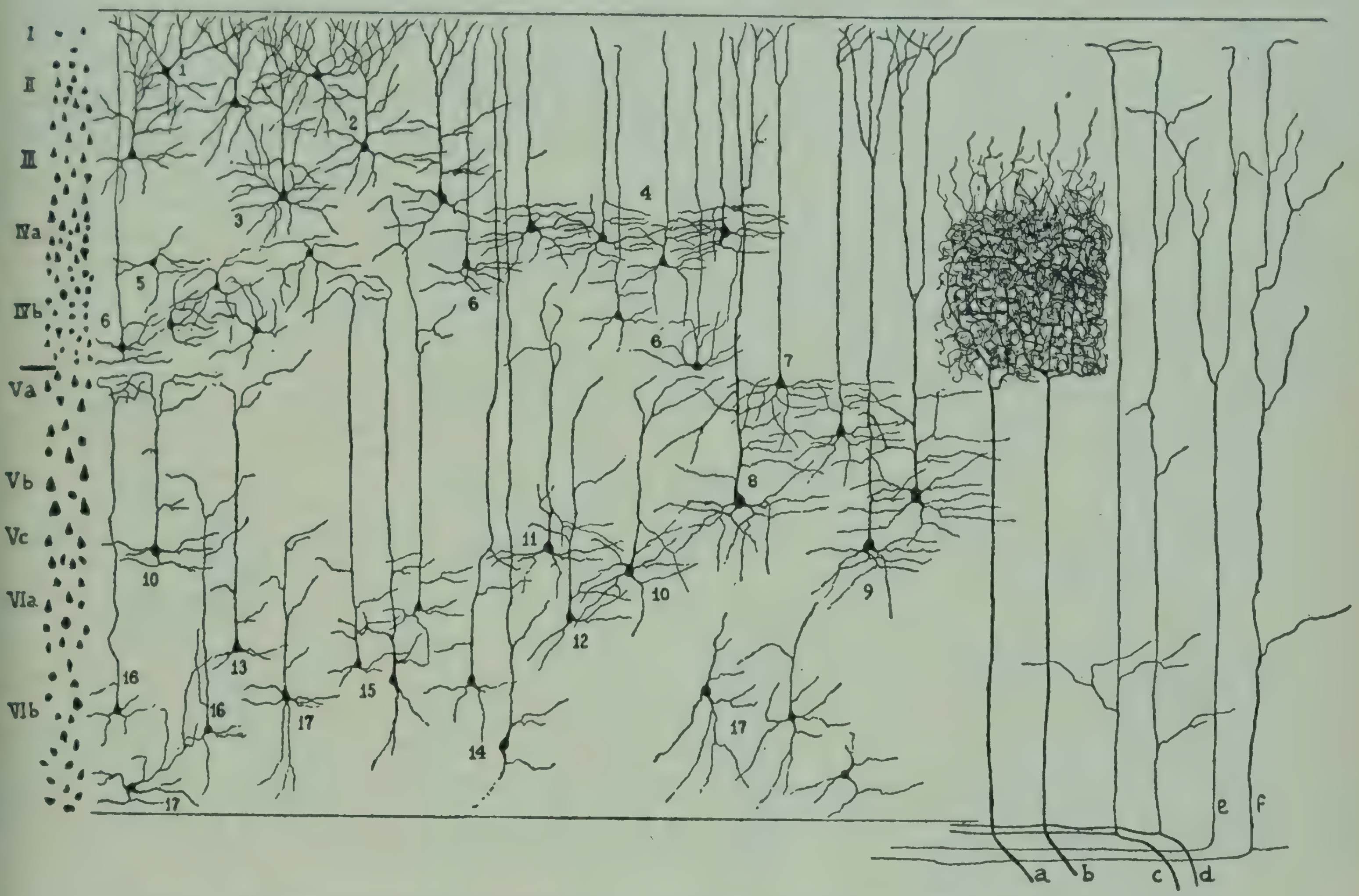


FIG. 205. Parietal Cortex of Mouse. On left a strip showing the cell layers. On right the principal afferent fibres, of which *a–d* are from thalamus, and *e, f* are association fibres. Main picture shows the relation of the chief types of descending axon cells to the cell layers, and the position of their dendritic branches. *For clarity of drawing the descending axons have been omitted*, but are shown in Fig. 203.

The cells are: 1, Pyramids of layer II.
2, 3, Pyramids of layer III.
4 to 6, Star cells and pyramids of layer IV.
7 to 11, Pyramids of layer V.
12, 13, Pyramids
14 to 16, Spindles } of layer VI.
17, Star cells

(LORENTE DE NÓ, in FULTON'S "Physiology of the Nervous System," Oxford University Press.)

Although this mapping of the cortex is valuable in stressing the varying structure of its different parts, it tells us little with regard to the functional connections of the neurones. For this purpose sections stained by Golgi's method are needed, and Figs. 203, 204 and 205 illustrate the findings thus made. It can then be realised that, as stressed by Lorente de Nó, the important point, as regards functions, is not so much the horizontal lamination as the fact that the cells are arranged in vertical chains, with synaptic connections at definite levels characteristic of each type of neurone. Seen from this standpoint we can recognise four main types of neurone in the cortex :—

(1) Cells with *axons that descend,* often into the white matter, where they form projection or association fibres (Fig. 203, 1 to 10, and Fig. 205).
(2) Cells with *axons that ascend* and ramify in other layers (Fig. 204, 18 to 21).
(3) Cells of Golgi type II, with short branching axons (Fig. 204, 23 to 26).
(4) Cells with horizontal axons (Fig. 204, 22).

It will be noticed that, with the exception of the projection and association neurones of the first group, all these neurones are intra-cortical, *i.e.* they must have internuncial functions.

Some typical neurones are shown in their proper positions in the layers, in Fig. 203 (descending axons), Fig. 204 (intra-cortical neurones, some with ascending axons) and Fig. 205 (general view of all types, and of the typical afferent fibres that reach the cortex).

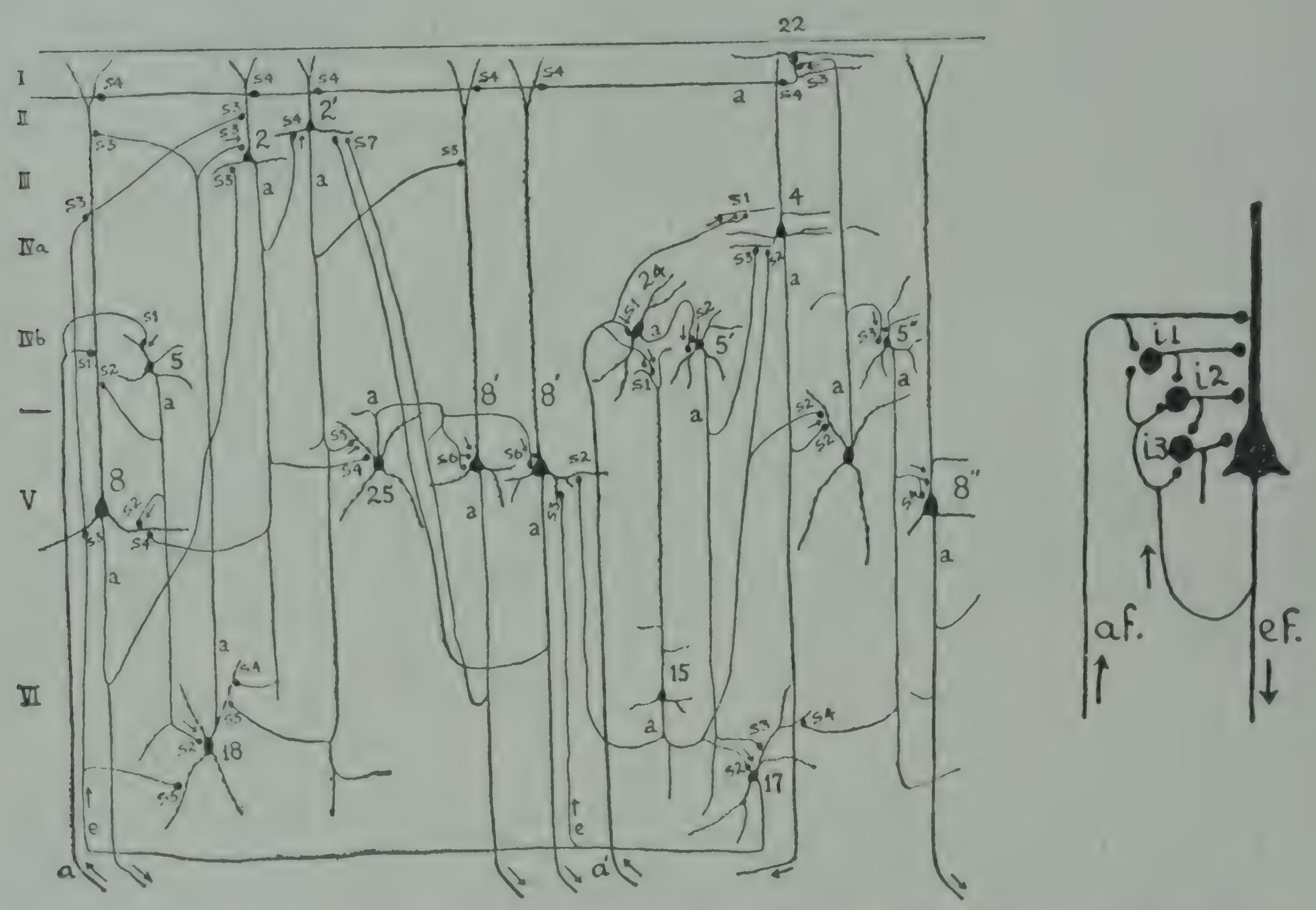


FIG. 206. Diagram of some of the Intracortical Chains of Neurones. The numbers on the cells, and the letters *a* and *e* on the fibres, are the same as in Figs. 203, 204 and 205. The axons are marked with *a*. Only a few dendrites and axonal branches have been included in the diagram. The synaptic junctions are indicated with the letter *s* (s_1, s_2, etc.) and with a thickening of the axon. It is assumed that the synapses marked with an arrow are passed by the impulses.

The small diagram at the right is a simplification of the diagram at the left. The afferent fibre *af.* activates the large pyramid, which is the origin of an efferent fibre *ef.*, and also a system of cortical internuncial cells (i_1, i_2, i_3) ; the recurrent collateral of *ef.* delivers impulses again to the internuncial system. This diagram summarises the plan upon which the cerebral cortex is built. (LORENTE DE NÓ, in FULTON's "Physiology of the Central Nervous System," Oxford University Press.)

In its general plan, as emphasised by Lorente de Nó,* the cerebral cortex appears to consist of chains of neurones, arranged in a generally vertical direction. Afferent fibres, from the thalamus or from other parts of the cortex, pass vertically upwards, and give off collaterals to all cell levels, except layers I and II. These collaterals form synaptic connections with all the types of cell, and we may speak of the afferent fibre, with all its connections as an *elementary cortical unit. In this unit the afferent impulse passes over to the*

---

* LORENTE DE NÓ. Article on Structure of Cerebral Cortex, in FULTON's "Physiology of the Nervous System," Oxford University Press, 1943.

efferent fibre. The general connections within an elementary unit, according to Lorente de Nó, are shown diagrammatically in Fig. 206.

The efferent fibres which emerge may pass, as association fibres, to other parts of the cortex, or as projection fibres to more distant parts.

A careful study of Fig. 206 and of the simplified diagram to the right of it, shows that a nervous impulse, entering the unit by an afferent fibre, makes synaptic contact not only with an efferent neurone, but also with numbers of internuncial neurones, by which reverberatory and delaying circuits are formed as by-passes, so that impulses are delivered over and over again to most of the cells, in the manner previously outlined. These trains of impulses not only result in more prolonged bombardment of the neurones, but, as a result of facilitation, may open up synapses not previously permeable.

So an entering impulse may, for instance, pass first to layer IV, thence down to layers V and VI and again up to layers II and III, finally emerging from a layer characteristic for the particular area.

It should be noted that all these arcs are superposed on the direct one from afferent to efferent, as *e.g.* from fibre *a* to cell 8 in Fig. 206. Further, there are various possible closed-chain circuits, as *e.g.* 8–2, 18–8, etc.; in the figure *a*, *a′* are afferent fibres, 8, 8′ and 8″ are effector neurones; the others (including 4 and 17, which give association axons *e*) are internuncial neurones. Some of the by-path arcs are confined to one layer, *e.g.* 25–8′, while others are long and span several different layers, *e.g.* 18–2. In the higher animals there are proportionally more of the short-axon connections, which confer greater delicacy and range of paths.

For Reference

Lashley and Clark (1946). The Cytoarchitecture of the Cerebral Cortex of Ateles: a Critical Examination of Architectonic Studies. *J. Comp. Neurol.*, **85**, 223.

LOCALISATION OF FUNCTION IN THE CORTEX

The demonstration by Fritsch and Hitzig in 1870 and by Ferrier in 1874 that stimulation of certain areas of the cortex led to rather accurately predictable movements on the opposite side of the body, and the discovery by Ferrier that fibres of the optic nerve ultimately ended at the occipital cortex, led to the conclusion that there is in the cerebral cortex a topographical differentiation of function. Subsequent detailed study over several decades confirmed this, and fitted in well with histological studies showing that the structure of the cortex likewise differed from place to place. Ultimately, it was expected, the cortex would be mapped out like a mosaic, each small area having one function, projicient or receptor, and one function only.

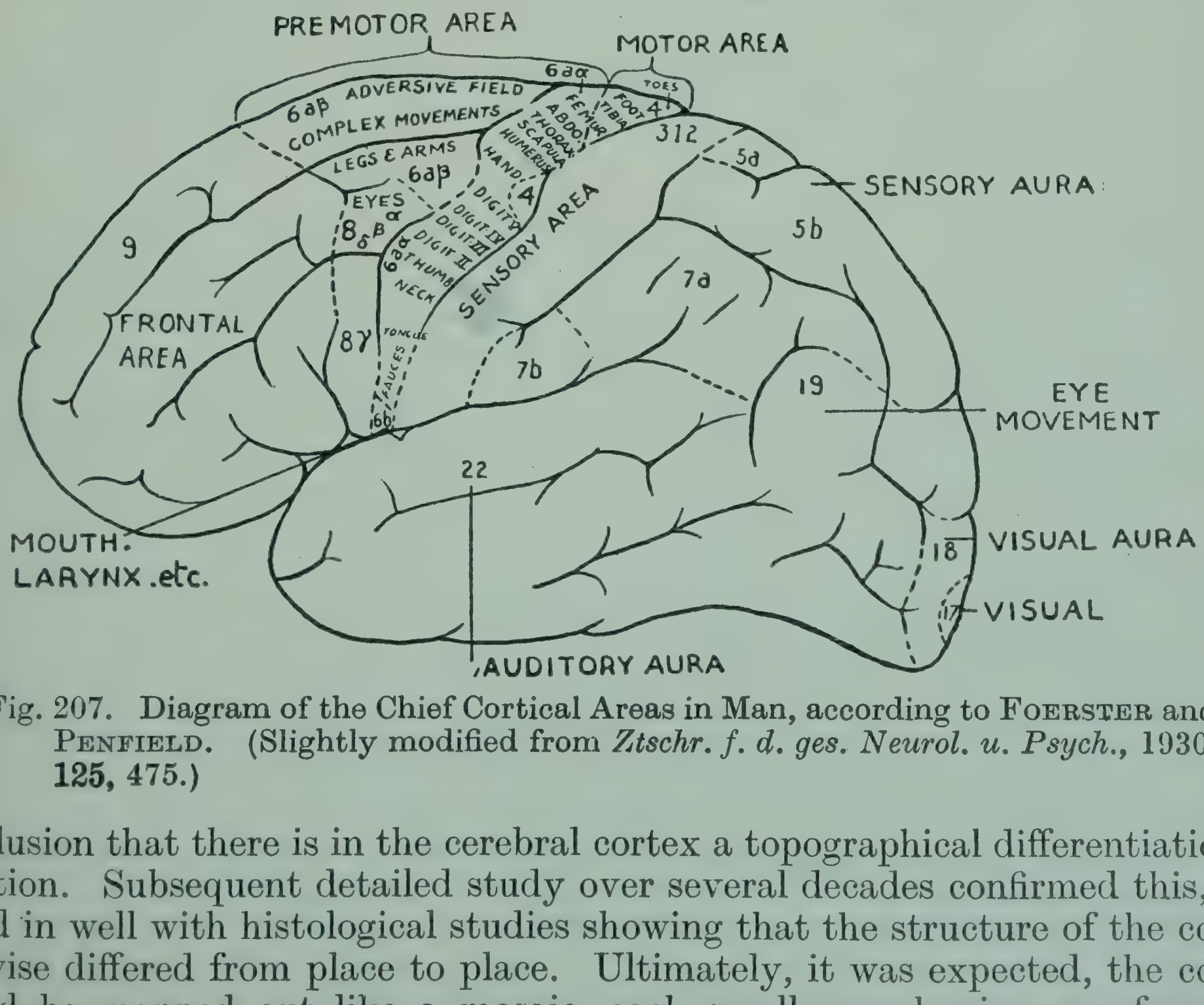

PREMOTOR AREA
MOTOR AREA
6aα
TOES
6aβ ADVERSIVE FIELD
COMPLEX MOVEMENTS
FEMUR
TIBIA
FOOT
4
ABDO.
THORAX
SCAPULA
HUMERUS
HAND
312
5a
SENSORY AURA
LEGS & ARMS
6aβ
EYES
α
8 δ β
4
DIGIT V
DIGIT IV
DIGIT III
DIGIT II
6aα
THUMB
NECK
SENSORY AREA
9
5b
7a
FRONTAL AREA
8γ
TONGUE
FAUCES
6b
7b
19
EYE MOVEMENT
22
MOUTH.
LARYNX .etc.
18
VISUAL AURA
17
VISUAL
AUDITORY AURA

Fig. 207. Diagram of the Chief Cortical Areas in Man, according to Foerster and Penfield. (Slightly modified from *Ztschr. f. d. ges. Neurol. u. Psych.*, 1930, **125**, 475.)

This conception was too simple and went too far; it accords neither with experimental facts nor with clinical observations. Localisation of a kind there certainly is, but known facts cannot be adequately explained unless it is realised that there is also considerable overlap. So far as the accepted motor areas are concerned, for instance, they are those which respond to threshold stimulation; the area mapped out as the motor area for the hand, for instance, is that which on the feeblest stimulation gives movements of the hand alone. It is in reality only predominantly a hand area, for on stronger excitation other movements result as well,* and these happen even when the cortical area is isolated from neighbouring areas. Moreover, on excision of a given motor area, movement in the part of the body it is supposed to govern is only transitorily abolished and is restored later, though perhaps with some clumsiness. Sensory areas similarly spread more widely than was formerly supposed.

Nevertheless, the accuracy of the localisations of function in the areas of lowest thresholds is undoubted.

Some areas are predominantly the site of origin of efferent projection fibres concerned with movements—the *motor areas*, while others, the *sensory areas*, are more intensively concerned with the reception and appreciation of sensory impulses. Others serve associative functions.

The functions and situation of these areas have been revealed by the study of the results of several different methods :—

(1) Anatomical and histological, by which it has been found that the different pathways to and from the cortex are connected with cortical areas that have characteristic structures. (2) The results of experimental or surgical removal, and of destructive disease, of various areas of the cortex. (3) The results of local electrical or chemical excitation (*e.g.* by application of strychnine) or of irritative disease, of the areas. (4) The *positive* electric potential changes in the cortex which result when sensory stimuli are applied to the appropriate receptors in deeply anæsthetised animals : under anæsthesia the sensory centres receive impulses in the usual way, but these do not reach the upper layers and so do not produce the widespread negative potential in the cortex which is the case in conscious animals, so the exact place of their arrival can be found.†

The functions of the chief of these areas of predominant function and their relation to the areas on the structural map are briefly as follows (Fig. 207, but not all areas are shown) :

The Frontal Lobe. Area 4 is subdivided into areas 4γ, 4a and 4s. Areas 4a and 4γ, the latter being the Betz cell area, are motor. In man, area 4γ occupies only the anterior lip of the Rolandic fissure. The axons of the Betz cells, and other cells of area 4, enter the pyramidal tract, and pass down, *viâ* interneurones, to the motor neurone pools in the brain stem and spinal cord.

Area 6, the so-called "*pre-motor*" *area*, occupies the greater part of the precentral gyrus. Its posterior part, adjoining area 4, is also motor, and sends fibres into the pyramidal tract. It also has other motor projections, but these are largely concerned with the extra-pyramidal system, and with the musculature of the mouth and respiration. Its lower parts (44 and 45) sometimes called Broca's area, are concerned with speech.

Area 8 is also largely extra-pyramidal and concerned with *eye movements*.

Area 9 *to* 12 is the prefrontal area or frontal association area, and is probably largely *associative* in function.

The Parietal Lobe. Area 3, 1, 2, the "post-central" area, is chiefly sensory, receiving most of the projections from the thalamus, for *cutaneous and general kinæsthetic sensitivity. Areas* 5 *and* 7 are also mainly sensory.

All these areas also contribute to the pyramidal tract.

The Temporal Lobe, Areas 41–42, receive the auditory radiations and are sensory for hearing.

Area 22 is an *auditory association* area.

---

* Walshe. *Brain*, 1943, **66**, 104; Murphy and Gellhorn. *Arch. Neurol. Psychiatr.*, 1945, **54**, 256.

† Adrian. *Lancet*, 1943, **245**, 33; *Brain*, 1943, **66**, 89.

The Occipital Lobe is entirely visual in function.

Area 17, the true *visuo-sensory* or "striate" area, receives projections from the lateral geniculate body and also sends fibres to it.

Area 18 is *visual associative.*

Area 19 is a visual area connected chiefly with the thalamus. Stimulation causes eye movements as a secondary effect.

Associative Functions of the Cortex. *The Pre-Frontal Area.* The earlier investigations on localisation led to a belief that it would ultimately be possible to allocate a precise function to every part of the cortex. This has proved not to be the case, and it must be admitted that large areas of the cortex do not possess such readily demonstrable functions, but might exercise general functions of some kind. Thus, large parts of the frontal lobes (areas 9 to 12) are quite unresponsive to electrical stimulation, and these areas on one side (notably on the right side) may be removed surgically without very evident effect. Removal of parts of the frontal lobes, or division of the thalamic connections of the lobes (frontal leucotomy) is an operation for relief of incurable anxiety states, and is often successful. Removal of the greater part of the areas 9 to 12 on *both* sides does, however, lead to general mental impairment and a change in personality.* Emotional reactions are changed, and the key to this change is perhaps in the important connections between the hypothalamus and the frontal cortex. Also, in children, disease of the pre-frontal area causes arrested mental development.

It has been suggested that these lobes play an important part in forming our habits of thought and behaviour and, in later life, in utilising the emotional drive to exertion and enterprise.† It seems, therefore, that these portions of the frontal lobes possess associative functions which are, however, probably shared by other areas of the cortex, so that they cannot, so to speak, be regarded as the seat of the "organ of mind." The association areas deal more or less as a whole with the complex processes that are interpolated between receipt of an impression and any consequent motor response.

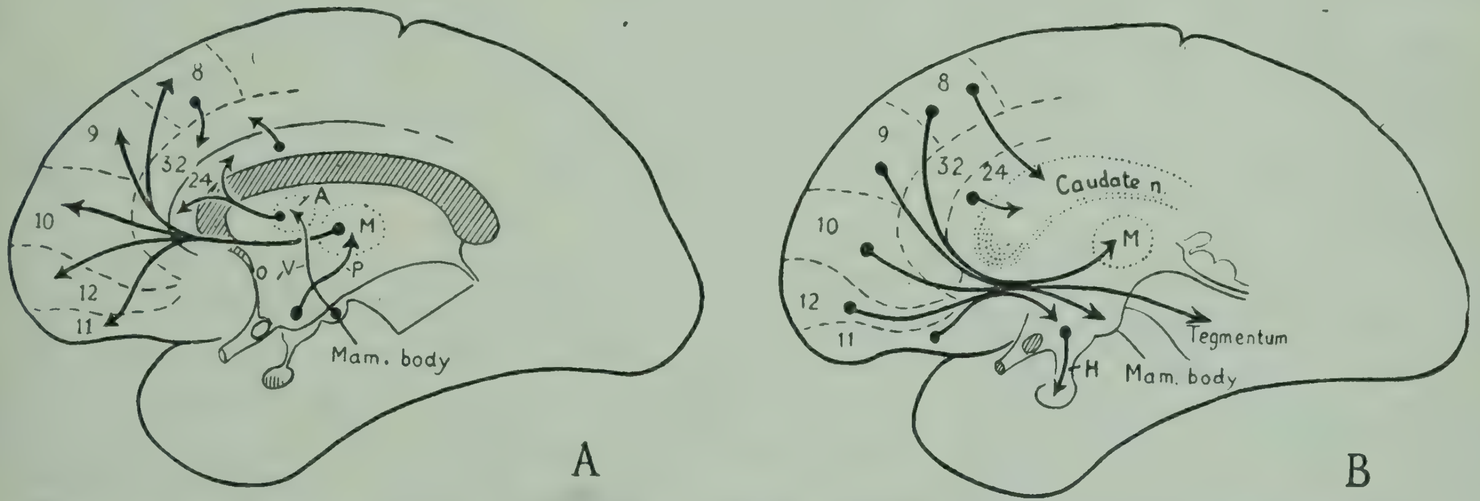


FIG. 208. Right cerebral hemisphere, medial aspect, showing (A) main afferent, and (B) main efferent connections of frontal lobes. Cortical areas are numbered. A, anterior, and M, dorsomedial nucleus of thalamus; P, fibres from hypothalamus; V, mamillothalamic tract (of Vicq D'Azyr); H, tract from hypothalamus to posterior lobe of pituitary. (LE GROS CLARK, *Lancet*, 1948, **254**, 353.)

Long association tracts passing backwards connect area 8 in the frontal lobe with area 18 in the occipital, and tracts also connect the frontal and temporal lobes. The frontal lobes are also receiving areas from various thalamic nuclei, of which some are relays from the hypothalamus: as shown in Fig. 208, these connections are both afferent and efferent.

It would seem that the frontal areas are important receiving areas from the hypothalamus, and thalamus; they are also efferent stations connected with various sub-cortical centres.

---

* JEFFERSON. *B. Med. J.*, 1937, (ii) 199.

† RITCHIE RUSSELL. *Lancet*, 1948, **254**, 356.

SENSORY AREAS IN THE CORTEX

The different types of afferent fibres that reach the cortex travel to definite areas according to their functions. It is curious that, for each locus of each sense there appears to be a double representation on either side: these are called, *e.g.* visual areas I and II, etc. The chief of these areas are in the parietal, temporal and occipital lobes.

To begin with general body or somatic sensations. These we have traced up from their origin on the opposite side of the body *via* the spinal, mesial and trigeminal lemnisci to the thalamus, at which level there appears to be a crude recognition of the particular quality of the sensation, coloured by an affective quality (*i.e.* a definitely painful or pleasurable property). Impulses are relayed on from the thalamus to the cortex of the parietal lobe, which they reach by passing through the internal capsule and corona radiata. The receiving area concerned is sometimes called the somæsthetic area.

The Somæsthetic Area. Its chief part includes the posterior lip of the fissure of Rolando and the greater part of the post-central gyrus lying posterior to it (areas 1 to 3). This post-central area has often been subjected to direct electrical stimulation in human subjects * with the result that sensations of numbness, pressure, touch or warmth were felt on the opposite side of the body, the actual place depending on the part of cortex stimulated. Thus, stimulation of the upper part of the area gave sensations in legs or feet; of the middle part sensations in trunk or arm, and of the lower part, sensations in face and head. More detailed examination has shown that the cortical representation closely follows the arrangement of the spinal dermatomes. Thus the order of sequence for the lower limb was hip, knee, ankle, toes (pre-axial surfaces), ankle, knee, hip, perinæum (post-axial aspects).† There was no special localisation of different types of sensation, but only representation of places stimulated.

According to some, sensation is also caused by stimulation of the corresponding parts of the anterior lip of the Rolandic fissure,‡ although, as we shall see later, this area (4 and 6) is also definitely motor. The explanation may be that what is represented here is proprioceptive sensibility from the muscles. Powerful excitation of the posterior parietal areas 5 and 7 may also produce sensation.

Another method of studying these areas in animals is by the local application of strychnine to the cortex.§ This causes an excitatory condition to spread out over the cortex from the place of application, and the animal responds by paying attention, *e.g.* by scratching, to the part of the body represented. The results are generally in agreement with those obtained on man. Yet another method is to record the potential changes on the surface of the brain when the skin is stimulated by any mechanical means, or when the dorsal roots are stimulated electrically, and to note from which regions characteristic waves can be recorded. This method shows that in addition to the areas in the upper part of the post-central gyrus (3, 1, 2) the dermatomal areas are also represented on the upper bank of the Sylvian fissure (area 7b). These areas are called somatic areas I and II respectively; they are reached by separate routes; somatic area II is also an extrapyramidal motor area.

Ablation of the areas concerned, or of parts of them, leads to loss of the finer grades of discrimination in sensation in the part represented ‖; but not to any complete loss of sensation. There is also clumsiness and inaccuracy in the movements of the parts concerned, and there is evidence that some of the cutaneous part of the sensations is bilaterally represented.

According to Head and Holmes the chief function of the somæsthetic area is the integration and critical examination of the crude sensory phenomena that are submitted from the thalamus. It thus has discriminative functions enabling us to form opinions regarding (1) relative *intensities* of sensations, *e.g.* of warmth ; (2) *spatial relationships, e.g.* the position of limbs, the exact locality of a skin stimulus ; (3) *similarities and differences* such as are arrived at by a synthesis of all the sensations involved. Thus we realise by touch similarities (or differences) of size, weight, texture. This is sometimes called *stereognosis*.

---

* Cushing. *Brain*, 1909, **32,** 44.

† Woolsey. *Feder. Proc.*, 1947, **6,** 437.

‡ Penfield and Boldrey. *Brain*, 1937, **60,** 389.

§ Dusser de Barenne. *Q. J. Exp. Physiol.*, 1916, **9,** 355; *Proc. R. S.*, 1924, **96B**, 272.

‖ Head and Holmes. *Brain*, 1911, **34,** 102.

The size of an area on the cortex is not proportional to the area of skin from which it receives impulses, but rather to its delicacy and importance : thus, in man the areas which represent hands, lips, tongue and face are relatively large, and those for leg and trunk small : in the pig, the snout, in horse the nostril, and in sheep the lips are the areas where representation is most minutely cared for, and accordingly these areas are relatively large.*

For Reference

BARD (1938). Studies on the Cortical Representation of Somatic Sensibility. *Harvey Lects.*, 1937–38, p. 143.

BUCY (1944). The Precentral Motor Cortex. (Univ. Illinois Press. Urbana, Ill.)

PENFIELD (1947). Some Observations on the Cerebral Cortex of Man. *Proc. R. S.*, **134**, 329.

WALSHE (1947). On the Contribution of Clinical Study to the Physiology of the Cerebral Motor Cortex. *Horsley Memorial Lecture.* (Edinburgh : Livingstone.)

The Auditory Centre. We have traced the fibres from the auditory nerve as far as the *medial geniculate body.* From this a new relay starts and, passing *viâ* the posterior limb of the internal capsule forms the *auditory radiation,* which ends in the auditory cortex in areas 41, 42. Not only is this sense served by the part of the cortex visible on the external surface of the temporal lobe, but also by cortex concealed from view (the gyrus of Heschl) in the depths of the Sylvian fissure. The cortex is divided physiologically into two parts, the audito-sensory area, and the audito-psychic area.†

The *audito-sensory* area (41 and 42) is largely hidden in the Sylvian fissure. A small part, however, is seen at the superior lip of the temporal lobe. Here the stimuli from the cochlea are interpreted as sensations of sound which vary in pitch, loudness and quality. The audito-sensory area is marked histologically by all its layers being very rich in both cells and fibres ; particularly is this the case with the outer layer of large pyramidal cells, which are both large and numerous. So also are the stellate cells in the layer just internal to this. These features are typical of the audito-sensory area. When sections are stained to show nerve fibres they exhibit a well-marked band running parallel to the external surface, *the line of Kaes.* The vestibular nerve has more important relations to the cerebellum, but it also has cortical connections with the temporal auditory area.

Electrical stimulation of the auditory area of one side in conscious human subjects gives rise to subjective noises in both ears ; and, because both sides are equally represented, a temporal lesion on one side, though it may cause some impairment of hearing, does not cause deafness.

There seems to be a spatial representation in the auditory cortex of the different parts of the cochlea. Thus, a low-pitched sound, which causes excitation of the receptors at the apex of the cochlea, or electrical stimulation of the fibres from that region, has been shown to produce electrical changes, signifying the arrival of nerve impulses, in the posterior part of the audito-sensory area.‡

The more peripheral portions of the auditory area in the lower two temporal gyri (area 22) form the *audito-psychic* area. Histologically, the cells found are small and there is no line of Kaes. This is an associative centre which is concerned chiefly with the appreciation of the significance of the sounds.

---

* ADRIAN. *Lancet*, 1943, **245**, 33.

† BREMER and DOW. *J. Neurophys.*, 1939, **2**, 308.

‡ WOOLSEY and WALZL. *Amer. J. Physiol.*, 1941, **133**, 498.

Stimulation causes auditory hallucinations and movements of the head to the opposite side. A special part of the superior and middle temporal convolutions on the left side (in right-handed people) is concerned with the understanding of speech, since failure of speech, or to understand spoken words, results from lesions in that situation (see Aphasia, Chapter XXVI).

The Visual Centre.* The cortical centre for vision is the "area striata" situated on the walls of the calcarine fissure (area 17). The visuo-sensory area on section shows to the unaided eye a conspicuous white striation known as the *line of Gennari* (fibre layer 4), which lies roughly midway between the two boundaries of the grey matter. It is formed by the interlacing dendrites of large stellate cells which replace the external layer of large pyramidal cells found in other parts of the cerebral cortex. This layer of large stellate cells is broad and is the most conspicuous feature of the visuo-sensory cortex. Other noteworthy features are the solitary giant pyramidal cells of Meynert, which are found forming a row of widely separated cells in the internal layer of large pyramidal cells.

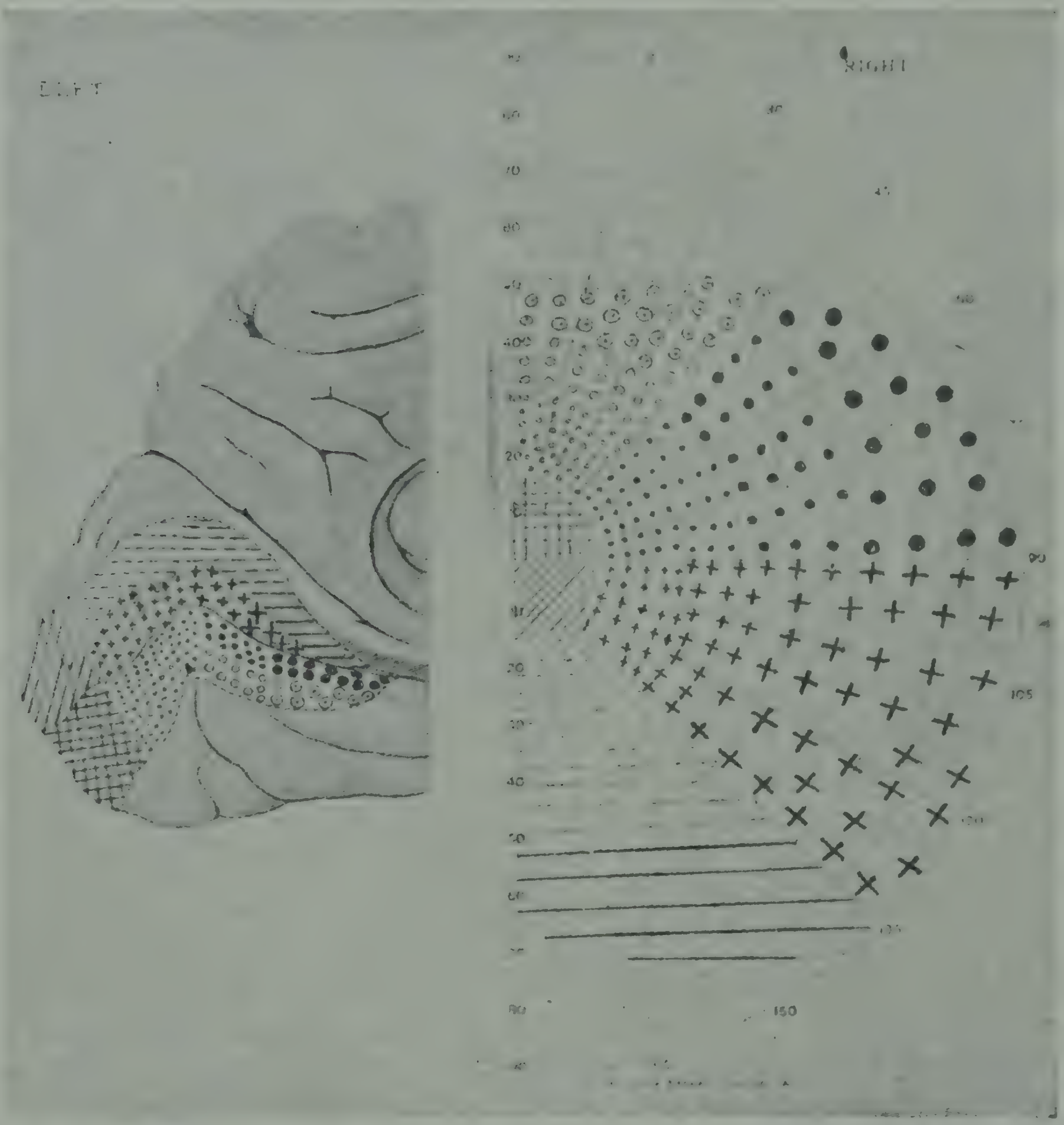


FIG. 209. The Relationship between the Field of Vision and the Calcarine Cortex. (After GORDON HOLMES.†)

To this area the fibres from the lateral geniculate body pass *viâ* the optic radiations (Fig. 198, p. 287) but exactly how they terminate is uncertain. The visual cortex on each side receives impulses from its own side of both retinæ, so that it is concerned, like most other cortical centres, with impressions arising from the opposite side of the surrounding world. There is fairly accurate spatial representation of retinal surfaces on the visual cortex (Fig. 209). Thus the upper quadrants of the peripheral retinæ are connected with the parts of the cortex above the calcarine fissure and the lower quadrants with parts below it. The macula is separately represented near the posterior occipital poles.

---

* GORDON HOLMES. *Proc. R.S.*, 1945, **132B**, 348.

† HOLMES and LISTER. *B. Med. J.*, 1919, ii, 193.

The histological delineation of the visual centre has been confirmed by the study of the electrical changes in the cortex when light falls on particular retinal localities in anæsthetised animals * and the projection of retinal areas on the cortical surface has been confirmed. It seems certain from these observations that when an image falls on the retina, a related image in the form of a pattern of electrical impulses is formed on the visual cortex.

Visual sensation is located in area 17, which is sometimes called the visuo-sensory area. Sensations of light and movements of the eyes result from stimulation of almost any part of the cortex surrounding the calcarine fissure. If the upper surface of the left calcarine zone be stimulated, both eyes move downwards and towards the right, as though light fell on the upper left quadrant of the retina from a source in the lower right quadrant of the visual field (Fig. 209). Excitation of the lower part causes movement of the eyes up and to the right for the converse reason ; while between these two parts there is an intermediate zone, most marked on the mesial surface, stimulation of which evokes a purely lateral deviation of the eyes to the left. Destruction of area 17 on one side causes *homonymous hemianopia*, *e.g.* if the right visual area is ablated there is loss of function of the right half of both retinæ which means blindness in the left half of the visual fields of both eyes.

The adjacent areas 18 and 19 and part of the posterior parietal area are concerned with the association of visual and other impulses. Here visual and tactile, and to a lesser extent auditory, impressions are correlated and synthesised to give proper visual perception and appreciation. A person born blind and having sight restored when adult is unable to understand the meaning of visual sensations until, by a process of re-education, these have been compared with the familiar tactile and auditory ones. Lesions in areas 18 and 19 in man result in disturbances of visual appreciation, so that written or printed words may lose their significance. Stimulation of area 19 causes visual hallucinations, and movements of both eyes to the opposite side.

Taste Centre. It now seems certain † that the cortical centre for taste is not, as long supposed, near to that for smell, but is represented on the cortical surface ; the nerves of taste (VII and IX) and those for somatic sensation from the face have a closely similar path. These nerves are fine and probably non-medullated. They pass to the nucleus of the *tractus solitarius*, from which relays pass by the mesial lemniscus to the posteroventral nucleus of the *thalamus*, and thence to the cortical area near to that receiving somatic sensory impressions from the face, *i.e.* about the lower end of the post-central gyrus.

Smell Centre. This is situated in the uncus and the anterior part of the hippocampal gyrus near to the auditory centre. Hence lesions in the temporal lobe often result in conjoint deficiency in hearing and smell. But the hippocampus is not entirely olfactory in function, and the central connections of this sense are still obscure.‡

MOTOR AREAS OF THE CORTEX

The areas of the cortex, the stimulation of which results in co-ordinated movements, are called *motor areas*. The most important ones are in area 4 and the posterior part of area 6, which are motor areas proper. Area 4 in man occupies the anterior lip of the fissure of Rolando, and area 6 most of the remainder of the precentral convolution. The areas extend on to the mesial surface of the hemisphere.

---

* Talbot and Marshall. *Amer. J. Ophthal.*, 1941, **24**, 1255.

† Börnstein. *Yale J. Biol. Med.*, 1940, **12**, 719 ; **13**, 133. Patton and Ruch, Howell's Physiology, 1946, 370.

‡ Brodal. *Brain*, 1947, **70**, 179.

Other Motor Areas. Movements can be excited by stimulation of various other cortical areas in addition to the precentral ones described above. The chief of these (Fig. 207) are :—

(1) Various areas in the lower part of area 6 cause movements of mouth, larynx, arrest of breathing, etc.
(2) The so-called "frontal eye-field" (area 8) causes conjugate deviation of eyes and head to the opposite side.
(3) Area 19 in the occipital region gives constriction of pupil, and, if strongly stimulated, conjugate deviation of eyes to opposite side.
(4) The post-central area (3–1–2) in the parietal lobe yields movements similar to those obtained from corresponding levels of area 4. These results in man are accompanied by sensations in the parts served, and are perhaps due to irradiation to area 4.
(5) Area 5 of parietal lobe and area 22 of temporal lobe may also cause adversive movements of head and eyes, also probably a secondary effect of sensory stimulation.

Of these areas, the frontal eye field is the most definitely related to voluntary movements, since ablation of that area leads to impairment of eye movements. The other areas are perhaps largely reflex in function, and concerned with postural reflexes, and such automatic acts as respiration, pupillary constriction, vocalisation, etc.

The following facts show that on cortical stimulation the seat of the excitation is in the grey, and not in the white, matter :—

(1) A smaller strength of current is required to excite the grey than the underlying white matter.

(2) Inexcitability of the grey matter can be produced by painting it with cocaine.

(3) The latent period between stimulation and movement is longer when the grey matter is excited than when a stimulus is applied to the white matter.

There is evidence that the stimulus acts directly on the pyramidal cells of layer V of the cortex, or on the endings of the afferent nerves to it.

Stimulation of Motor Centres. The earlier stimulation experiments were made on the lower mammals, in which the anatomical relations are somewhat different and need not be considered here, but valuable information was obtained from them, and was extended and applied in experiments on monkeys and apes.* Ultimately many experiments have been carried out on the human subject,† and these have confirmed the inferences drawn from the work done on the apes. In man the localisation, though not punctate in precision, is fairly well defined, especially in area 4.

If a part of the precentral convolution be stimulated with a weak faradic current, movements are produced on the *opposite* half of the body, the site of the movements depending on the locality of the spot stimulated.

We may say broadly that, by stimulation of the precentral convolution at a series of points from above downwards, we get movements in that order of the leg, arm and face ; though, as is shown in the diagram (Fig. 207), within these larger areas smaller areas can be distinguished for definite co-ordinated movements of the smaller parts of the body. Those portions of the limbs, such as the fingers and thumbs, which execute delicate movements have larger areas of representation in the cortex than those whose movements are less precise. This means that the number of muscle fibres in a motor unit is smaller when the muscle is capable of delicately adjusted movements, and hence a larger cortical area, with more motor pyramidal cells, is needed to innervate it and its associates.

Representation is not of muscles, but of movements, *i.e.*, the movements obtained by excitation of the motor areas resemble in every respect the co-ordinated movements observed during the normal activity of the animal, and involve the reciprocal innervation of antagonistic muscles. The movements are organised by co-ordination of muscular contractions in both time and space. Never do we find simultaneous contractions of antagonists, even where two opposing centres are excited simultaneously ; one reaction is prepotent, as is the case with reflexes.

---

* Grünbaum and Sherrington. *Proc. R. S.*, 1901, **69B**, 206 ; 1903, **72B**, 152 ; Leyton and Sherrington. *Q. J. Exp. Physiol.*, 1917, **11**, 135.

† Bidwell and Sherrington. *Brit. Med. J.*, 1893, (ii) 988 ; Penfield and Boldrey. *Brain*, 1927, **60**, 389.

Thus the movement of opening the jaw, excited from the cortex, involves a relaxation of the normal tone of the masseter muscle. Flexion of the leg demands relaxation of the extensor muscles. As in the spinal reflexes, this relaxation can be abolished under the systemic action of strychnine or of tetanus toxin. Excitation of the cortical centre for the movements of the jaw then causes contraction of both closers and openers of the jaw, *i.e.* the stronger masseter muscles predominate, so that the jaw is firmly closed.

The part played by muscular relaxation in the response to cortical stimulation is also well seen in the eye muscles. Stimulation of the centre for eye movements on area 8 of the frontal lobes on the right side causes " conjugate deviation " of both eyes to the left. This movement involves contraction of the right internal rectus and left external rectus and a simultaneous inhibition of the tone of the right external rectus and left internal rectus. If all the nerves of the right eye be divided except the VIth, which supplies the external rectus, this eye looks permanently towards the right side, *i.e.* a right external strabismus or squint is produced. On now exciting the right cortex, *both* eyes still move towards the left, although the right internal rectus is paralysed. The movement of the right eye stops at the middle line, however, and is brought about simply by a relaxation of the tone of the right external rectus muscle. (Sherrington.)

This movement of both eyes on stimulation of one side of the brain shows that the function of each hemisphere is not entirely unilateral with regard to the muscles of the body. As a rule, the response to excitation of the motor area for limbs is only contralateral. In the case of those movements, however, which are normally carried out by co-operation of the muscles of the two sides, such as the movements of the trunk, neck, and eyes, stimulation of the motor area in one hemisphere evokes a movement involving the muscles of both sides of the body, *i.e.* the cortical representation is one of movement rather than one of muscles. Also, where an action is carried out by similar contractions of corresponding muscles on the two sides, the movement itself is bilaterally represented in the cortex. Types of such reactions are found in closure of the mouth, contractions of the abdominal muscles, extension or flexion of the trunk. It seems that under such circumstances there is a free communication between the lower motor centres of the two sides, since the bilaterality of the response is not altered by extirpation of the cortex of the hemisphere opposite to that which is being stimulated.

A certain antagonism is evident between the action of the cerebral hemispheres and that of the lower centres. Whereas in the decerebrate animal there is increased tonus in the extensor muscles, stimulation of the cortex produces flexion more easily than any other movements.

The *kind* of response to excitation is never predictable ; the same point may at one moment give flexion and at another extension. Moreover the response can vary according to the immediately preceding history of the area stimulated. Sub-threshold stimuli may be facilitated and so rendered adequate by preceding weak stimuli.* When a point is repeatedly stimulated for short periods, the resultant movement becomes larger and larger (" primary facilitation ") ; moreover, the area of cortex from which that movement can be evoked grows in size, and spreads into neighbouring areas that previously gave other responses. This is called the " deviation of the response," and is due to an irradiative facilitation (" secondary facilitation "). It probably means that a given movement, although represented by large numbers of pyramidal cells in the usual " area " of representation, is also represented by cells in other areas, *i.e.* there is considerable overlap. Such deviated responses may occupy large areas of the cortex.†

Under some conditions, cortical stimulation is followed by a period of *subnormal* excitability. This is called " extinction."

Stimulation of points in area 6 in man gives movements in the same parts as is obtained by excitation of the adjacent portions of area 4, but stronger stimuli are needed, especially at the rather diffuse anterior edge of the motor area. Stimulation of area 6 greatly facilitates the responses from area 4 ; it seems likely, therefore, that the cells responsible for motor representation become more sparse as we pass forwards in area 6, and finally vanish along an indefinite boundary line. If the pyramidal tract has been divided, or area 4 destroyed, it is claimed that strong excitation of area 6 still produces general adversive movements, *i.e.* rotation of head, eyes and trunk to opposite side.

---

* Graham Brown. *Q. J. Exp. Physiol.*, 1916, **10**, 103.

† Graham Brown and Sherrington. *Proc. R. S.*, 1912, **85B**, 250.

It is claimed by Fulton, Foerster and others that area 6, which they call the *pre-motor area*, is not only connected *viâ* area 4 with the major projections of the pyramidal tracts, but also has other connections, the so-called extra-pyramidal paths, with the spinal motor neurones. Others, *e.g.* Walshe,* regard area 6 as a region for synthesis of movements executed *viâ* the pyramidal system, and therefore as a higher level of motor control. The movements are more precise and more nearly confined to one muscle the nearer the stimulated point lies to the central fissure, *i.e.* the closer to area 4. As area 6 is approached the movement tends to become that of a whole limb.

Cortical Epilepsy. When electrical excitation at any strength above the minimum effective stimulus is applied to the motor area of the cortex, the movements evoked tend to persist for a short time beyond the duration of the stimulus. This is readily understood to be due to the same causes as produce after-discharge in spinal reflexes, viz. deviation of impulse along delay-circuits. On still further increasing the strength of the current, the contraction spreads to adjoining muscles, and finally may affect all parts of the body, giving rise to an epileptiform convulsion. The same effect may often be caused by weak stimuli if the irritability of the cortex be raised in consequence of previous repeated stimulation. The fit consists of two parts. The first effect of the stimulation is a strong tonic contraction ; this outlasts the stimulus for some time, and then gives way to a series of clonic contractions, repeated at first at intervals of from six to ten per second, but gradually getting slower as the fit dies away.

The main phenomena of a "convulsion beginning unilaterally," due to irritation of any portion of the motor area, were described by Hughlings Jackson in 1864, before the experimental proof of cortical localisation had been brought forward by Fritsch and Hitzig. They may occur in the human subject as a result of irritative lesions of this part of the cortex, such as that due to the presence of a tumour or a spicule of bone pressing on the brain. They commence in one of three predominant foci, viz. thumb, angle of mouth, or hallux. Jackson showed that in this condition the convulsive movements follow a certain order or "march." Thus, if the thumb area be the seat of stimulation, the fit begins by a contraction of the thumb muscles, then spreads to the muscles of the hand, fore-arm and shoulder of the same side, and then to the face, trunk and leg. If it begins in the toe, the order would be up the leg and trunk and down the arm. The same "march" is observed in artificial stimulation of the motor area, and the explanation is similar in both cases, viz. the excitation spreads by the overlapping of cortical fields, and "deviation" follows in a predictable sequence.

Suppressor Strips. Certain areas of cortex on excitation lead not to muscular movement, but to the *suppression* of any contralateral movement. The chief of these is called area 4s, which is a narrow band between areas 4 and 6. Other suppressor strips are 2s, 8s and 19s. These are instances of the complex relations between one part of the cortex and another. The mode of action of the strips is discussed later.

Ablation of Motor Areas. The results of ablation of the motor areas are more serious in man and the higher apes than in lower mammals. In the cat or dog, for instance, removal of the motor areas on one side results in nothing more than a transitory paralysis and spasticity of the musculature on the opposite side of the body.† The effect is greater and lasts longer on the fore limbs, but in a few days the paralysis nearly disappears and only a little weakness and loss of skilled movements remains. The spasticity is an extensor rigidity and lasts somewhat longer than the paralysis. Bilateral removal of the accredited motor areas is also not followed by lasting paralysis and we may take this to mean that the more diffuse and less excitable motor centres are brought into function.

In the higher apes ‡ and in man § the results are more serious and long-lasting. Thus, if the left arm area (4b) is removed from a chimpanzee, after having its situation verified by electrical stimulation, there is at first loss of reflexes, and complete and flaccid paralysis of the right arm. After a period which varies from days to weeks there is a progressive partial return of function from the shoulder downwards, the hand never recovering its finest movements completely. If the arm area on the right side is removed at this stage no effect is seen on the left arm. These observations have been interpreted by some as showing that there is some other motor centre and some path other than the pyramidal tract by which movements are controlled. Opinion is not unanimous on this, however.

---

* Walshe. *Brain*, 1935, **58**, 49.

† Laughton. *Amer. J. Physiol.*, 1928, **85**, 78.

‡ Ferrier. *Phil. trans.*, 1875, **165**, 433 ; Leyton and Sherrington. *Q. J. Exp. Physiol.*, 1917, **11**, 135 ; Grünbaum and Sherrington. *Proc. R. S.*, 1903, **72B**, 152.

§ Foerster. *Brain*, 1936, **59**, 135.

When in man, or one of the apes, the whole of the precentral motor area (*i.e.* areas 4 and 6) are removed on one side, the affected limbs on the opposite side are paralysed, and their reflexes are at first depressed; in a few days, however, spasticity may develop and extensor reflexes become exaggerated.[*] Voluntary power returns partially in a few days, but is lost again, and permanently, if the motor areas of the opposite side are removed. The posture of the animal then resembles that of the decorticate or thalamic preparation. The effects of removing area 6 (the "premotor" area) alone is also stated to be to cause spasticity and exaggerated extensor reflexes, with involuntary grasping movements of the hands.

These results are capable of at least two explanations. According to Walshe[†] the precentral motor areas comprise both area 4 and area 6, of which area 6 represents a higher level of function, largely associative in character. Spasticity only develops when the pyramidal tract is almost or completely severed from the whole area. Fulton, on the other hand, regards area 4 as the principal area of origin of the pyramidal tract, and area 6 as not only contributing to the pyramidal tract directly, and also indirectly by association with area 4, but also regards both areas, but chiefly area 6, as having other important connections with lower centres. These are called the extra-pyramidal paths and will be referred to later.

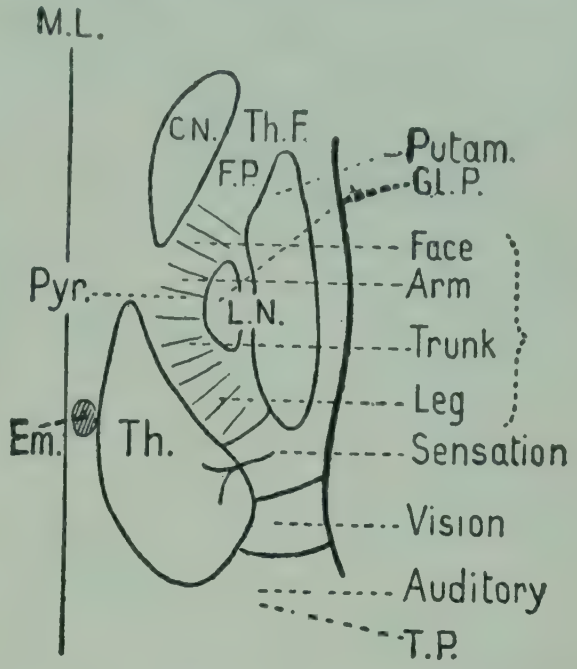


FIG. 210. Diagrammatic Horizontal Section through Base of Brain showing Relation of Internal Capsule to the Basal Ganglia. Anterior end is above.

ML = Midline; Pyr. = Pyramidal tract; Em = Supposed path connected with emotional expression; C.N. = Caudate nucleus; Th = Thalamus; Th. F. = Thalamo-frontal fibres; F.P. = Fronto-pontine fibres; Putam = Putamen; Gl. P. = Globus Pallidus; L.N. = Lenticular nucleus; T.P. = Temporo-pontine fibres. Other words indicate position of fibres for movement of parts or fibres for sensations indicated. These fibres mostly pass vertically to plane of paper. (SAMSON WRIGHT. "Applied Physiology," Oxford University Press.)

EFFERENT (OR CORTICO-FUGAL) PATHS

Two types of projections leave the cortex to proceed to lower centres, viz. the pyramidal and allied tracts which proceed to motoneurone pools in the bulb or spinal cord; and "extra-pyramidal" projections, which pass in the first instance to sub-cortical stations.

I. **The Pyramidal or Cortico-spinal Tract.** As we have seen (p. 226), this tract is of mixed origin, from areas 4 and 6, from the parietal lobe, and possibly from sub-cortical regions.

The "upper neurone" fibres of the pyramidal tract pass by the corona radiata through the internal capsule, mid-brain, pons, medulla and spinal cord, to end mainly in relation to contralateral dorsal horn cells, from which impulses pass *viâ* intermediate column cells to motoneurones of the ventral horn of the cord.

The *internal capsule* is a band of fibres lying between the lenticular nucleus laterally and the thalamus and caudate nucleus medially. The position of the various motor and sensory fibres in the internal capsule is shown in Fig. 210. The motor fibres occupy the genu and the anterior two-thirds of the posterior limb, the fibres in order from before backwards being eyes, head, shoulder, arm, hand, trunk, leg, foot.

---

[*] FULTON. *Brain*, 1935, **58**, 311.

[†] WALSHE. *Brain*, 1935, **58**, 49.

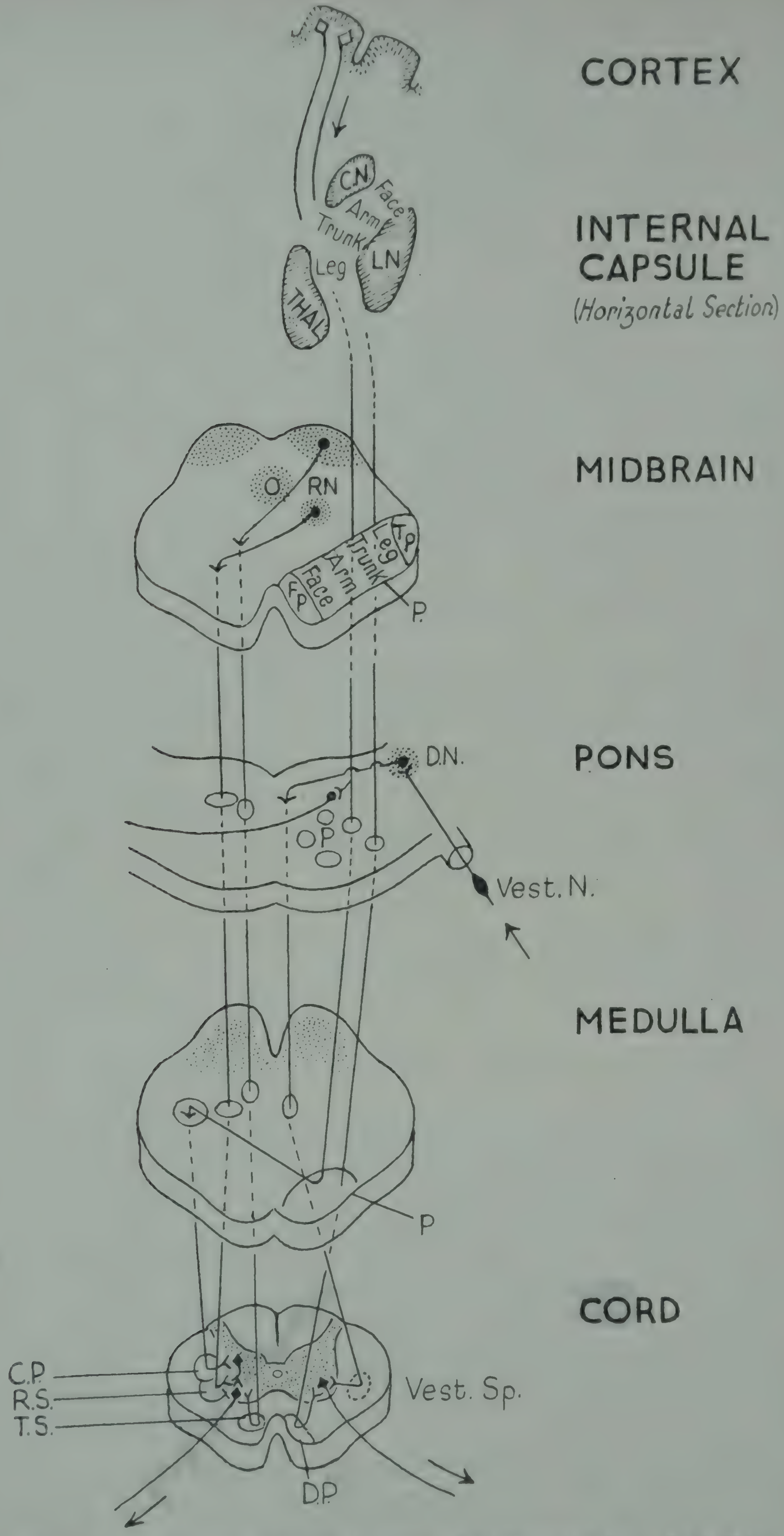

CORTEX
CN
Face
Arm
Trunk
Leg
LN
THAL
INTERNAL CAPSULE
(Horizontal Section)
MIDBRAIN
RN
Tp
Leg
Trunk
Arm
Face
Fp
P.
PONS
D.N.
P
Vest. N.
MEDULLA
P
CORD
C.P.
R.S.
T.S.
Vest. Sp.
D.P.

Fig. 211. Diagram of Chief Efferent Paths arising from Left Side. CN, Caudate nucleus; LN, Lenticular nucleus; THAL, Thalamus; RN, Red nucleus; Tp, Temporo-pontine tract; Fp, Fronto-pontine tract; P, pyramid; D.N., Deiters' nucleus; Vest.N, Vestibular nerve; C.P., Crossed pyramidal tract; R.S., Rubro-spinal tract; T.S., Tecto-spinal tract; D.P., Direct pyramidal tract; Vest. Sp., Vestibulo-spinal tract.

The motor fibres pass on as a compact mass (Fig. 211), forming the middle portion of the descending main pathway which passes through the crusta of the mid-brain ventral to the substantia nigra. Retaining the same situation, they reach the pons. They are here split up into discrete bundles by the cross passage of the pontine fibres. On leaving the pons and entering the medulla they form into two bundles once more. At the lower part of the medulla the main mass of fibres decussates to reach a more lateral position (the motor decussation), forming the lateral cortico-spinal or crossed pyramidal tract (about 10 per cent. of its fibres are not crossed however). A remnant of fibres, which is called the anterior cortico-spinal, or direct pyramidal tract, is left behind, close to the anterior fissure. Both direct and crossed pyramidal tracts descend in the cord in these two situations, turning in at different levels to relay at dorsal horn cells, the crossed ones to cells of the same side, the uncrossed ones to cells of the opposite side. In this way, fibres that have originated in the motor area on the one side of the brain ultimately innervate anterior horn cells which send fibres to muscles on the other side of the body.

Although the pyramidal tract is regarded as the principal motor path, section of the tracts at the level of the pyramids does not result in complete motor incompetence in lower animals. In the cat, for instance, Liddell found automatic movements normal, but skilled movements, *e.g.* walking along a ladder, impaired. If the motor cortex is also removed skilled movements are quite impossible. In man (*v.* p. 233) the effects of pyramidal section are more serious and complex.

Various collateral paths are formed from the pyramidal tracts. The chief of these are to the nuclei pontis, from which fibres pass in the middle peduncle to reach the cerebellum on the opposite side. Other branches of the tract cross in the mid-brain, pons and medulla to reach the motor nuclei of cranial nerves on the opposite side.

II. **The Extra-pyramidal System.** The evidence obtained from ablations of the precentral cortex in man and the higher apes, together with some clinical evidence, points to there being other efferent paths besides the pyramidal tracts by which the cortex is connected with lower motor centres. Thus, the eye movements controlled from area 8 in the second frontal convolution, and the complex movements of body and limbs obtained from area 6 of the precentral area, are obtainable after section of the pyramidal tracts. These so-called extra-pyramidal tracts still remain somewhat obscure, but those for the trunk and limbs appear to have important connections with the basal ganglia. The chief extra-pyramidal projections, some of them non-medullated, are to :—

(1) *The reticular formation* of the brain stem, from all four lobes, but chiefly from areas 4, 6, and 3, 1, 2.

(2) *The substantia nigra, red nucleus* and other grey matter of the brain stem.

(3) *The corpus striatum*, especially from the cortical suppressor bands to the caudate nucleus.

(4) *The thalamus* from all four lobes.

(5) *The hypothalamus.*

Some will relay at more than one of these stations.

The simplest of them is the path from the frontal eye field, called the *cortico-nuclear* tract. The fibres pass in the anterior limb of the internal capsule, with the fronto-pontine fibres, thence in the mesial fifth of the crusta of the mid-brain, from which fibres pass dorsally to the oculomotor nuclei of the opposite side. There is little difference, beyond anatomical detail, between this path, or paths to bulbar motor nuclei, and the pyramidal path.

The other cortical projections have a more complicated pathway and function. They arise from the motor areas, which all include some extra-pyramidal projections, in addition to pyramidal ones, *e.g.* from the precentral areas, and especially from area 6.

Thus after section of the pyramidal tracts, a flaccid paralysis results, and stimulation of area 4 then gives movement, or, more usually, an inhibition of any postural tonus.[*] Impulses passing along the extra-pyramidal projections seem especially to exert a modifying influence, which may be either inhibitory or facilitating, upon the lower motor neurones, rather than themselves to excite those neurones. This may be demonstrated by eliciting a movement, either reflex or by cortical excitation, at regular intervals, and then studying the effect on that movement of applying stimuli to the reticular formation of the brain stem, or other cell stations on the extra-pyramidal path. As shown in Fig. 212, the result of such stimulation is to abolish, or to enhance, the movement, according to the region stimulated (Fig. 213). The most common result is inhibition, but by hypothalamic stimulation facilitation can be obtained.[†]

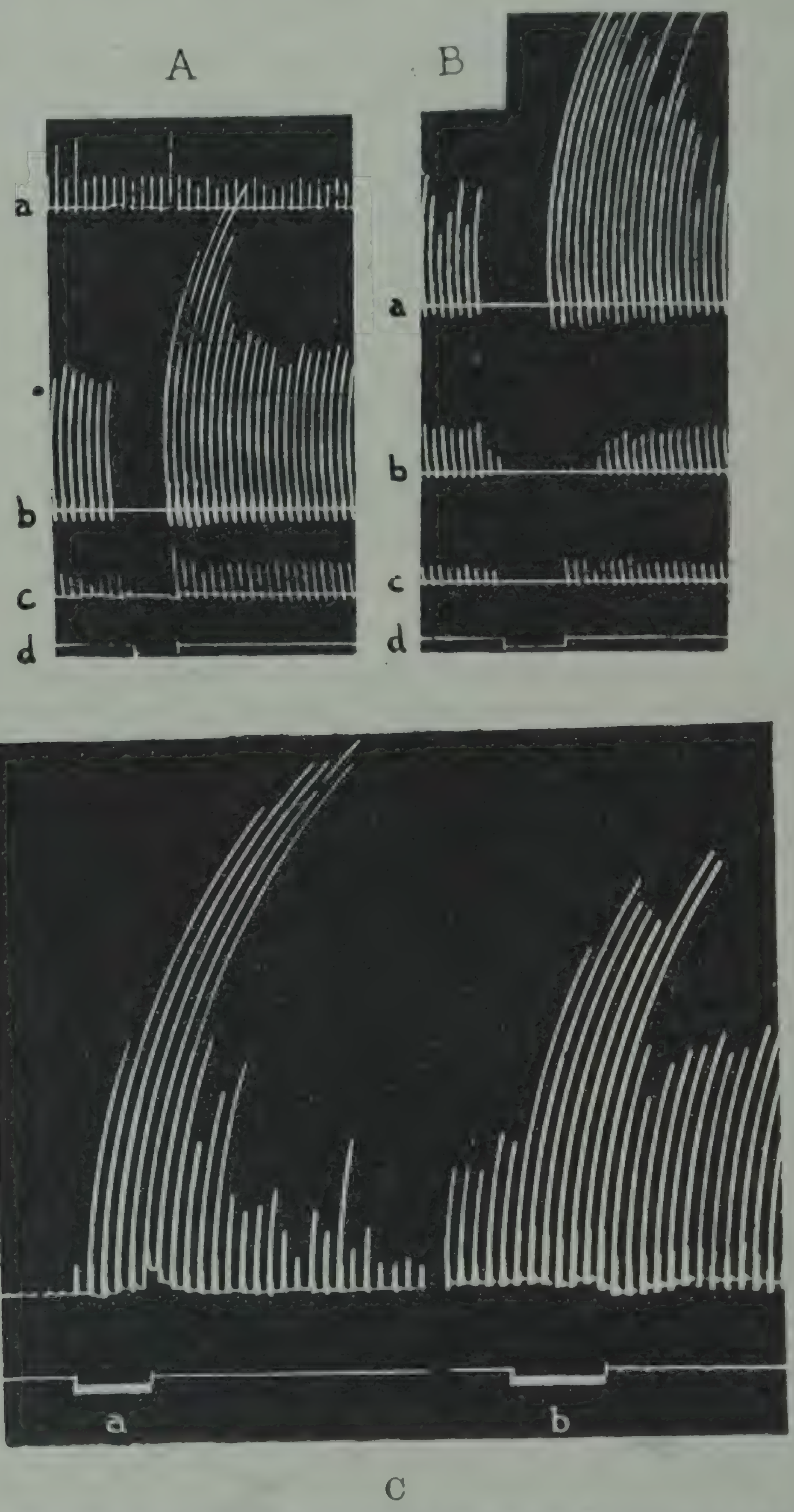


Fig. 212. A, B, Effects of stimulation of reticular formation of brain stem on flexor reflex (a), knee jerk (b), and blink reflex (c) evoked at 2 sec. intervals. Bulbar stimulation at line (d). C, Effects of stimulation of hypothalamus on response to cortical motor response (a) and patellar reflex (b). The experiments were on monkeys. (Magoun. *Quart. Bull. N.U.M.S.*, 1946, **20,** 180.)

It is probable that the extra-pyramidal path is interrupted by one or more relays, and finally ends at the motor cells of the lower motor neurone, as diagrammatically suggested in Fig. 213. The reticulo-spinal, tecto-spinal and rubro-spinal tracts are examples of large bundles of fibres having extra-pyramidal functions. The reticulo-spinal tracts convey both facilitatory and inhibitory impulses, the former lying more dorsally, and arising from cells higher up in the brain stem than the latter.* The suppressor strips and area 6 are important cortical areas from which extra-pyramidal paths originate, and some of these fibres have relays in the reticular formation, etc., and are inhibitory in action; but probably most other cortical areas contribute to the system. Damage to area 6, according to this view would result in spasticity by leaving the lower motor neurone in a state of release.

---

[*] Marshall. *Arch. Neurol. Psych.*, 1934, **32,** 778; Tower and Hines. *Science*, 1935, **82,** 376.

[†] Magoun and Rhines. *J. Neurophysiol.*, 1946, **9,** 165.

Kinnier Wilson † pointed out that destruction of the corpus striatum in man may give rise to a condition (Parkinson's disease), characterised by tremor and weakness in the execution of willed movements and increased tonicity of the muscles. He therefore ascribed to these bodies, or rather to the sensori-motor mechanism which has its chief meeting-place in their nuclei of grey matter, a steadying effect on the motor system.

The condition of hypokinesia, or weak movements can be reproduced experimentally in monkeys by lesions in the globus pallidus or hypothalamus; we may suppose then that in Parkinson's disease it is the facilitating pathway that has been eliminated, leaving the lower motor neurone unduly inhibited.

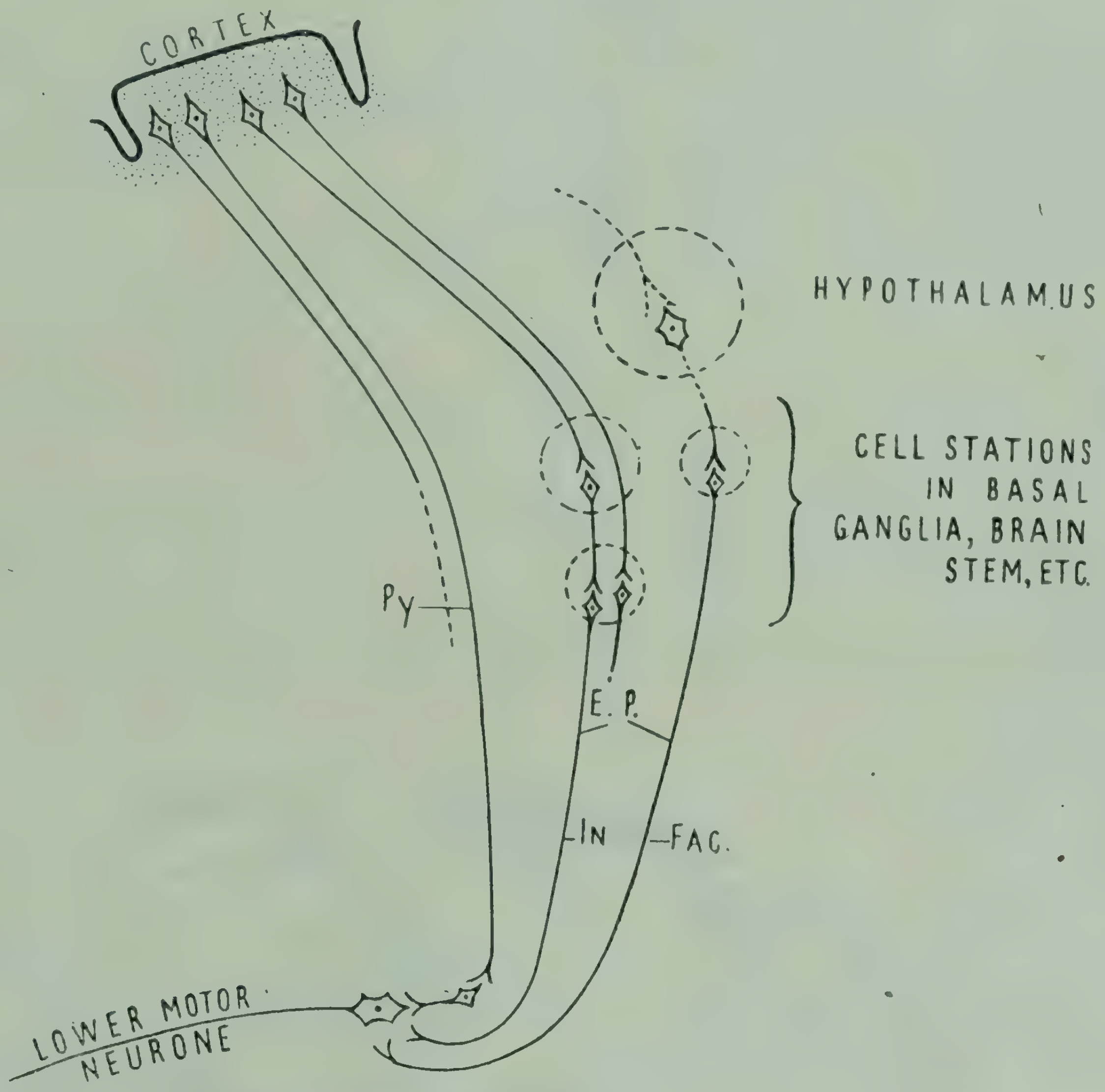


FIG. 213. Diagram of Pyramidal and Extra-pyramidal projection systems. *Py*, pyramidal; *E.P.*, extra-pyramidal. *In*, path producing inhibition of lower motor neurone; *Fac*, path producing facilitation.

It will be seen that there are important connections with the basal ganglia.

Functions of the Basal Ganglia. The basal ganglia include:

(1) The corpus striatum.
    (i) Caudate nucleus.
    (ii) Lenticular nucleus.
        (*a*) The putamen.
        (*b*) The globus pallidus.
(2) The claustrum.
(3) The amygdaloid nucleus.
(4) The body of Luys.

The relations of the lenticular nucleus are shown in Figs. 210 and 215.

The basal ganglia receive important afferent connections from the cortex, but these were missed for a long time because they are non-medullated.‡

---

* NIEMER and MAGOUN. *J. Comp. Neurol.*, 1947, **87**, 367.
† WILSON. *Arch. Neurol. Psych.*, 1924, **11**, 385.
‡ GLEES. *J. Anat.*, 1944, **78**, 47.

Attempts to produce effects by stimulation also failed, until it was realised that the results of stimulation were inhibitory and not motor. If any movement due to cortical action is in course, *e.g.* as a result of stimulating area 4, stimulation of the caudate nucleus or globus pallidus will at once arrest it. As has been stated above, such a result can be got by the action of extra-pyramidal systems which can inhibit the lower motor neurone, and part of the effect is to be explained in that way. But the inhibitory action is also one upon the cerebral cortex itself. Light was thrown on the functions of the basal ganglia when, in studying the action of the suppressor strips already referred to, Dusser de Barenne and McCulloch * followed the action potential which resulted when the strip, *e.g.* area 4s, was stimulated, *e.g.* by local strychnine application. They found that as a result of this the caudate nucleus became active; this then passed on impulses to the globus pallidus and this again to the thalamus. From the thalamus, the impulses passed back to the related motor area, *e.g.* area 4, which was thereby inhibited (Fig. 214). In this manner the basal ganglia could play an important part in regulating and controlling muscular responses,

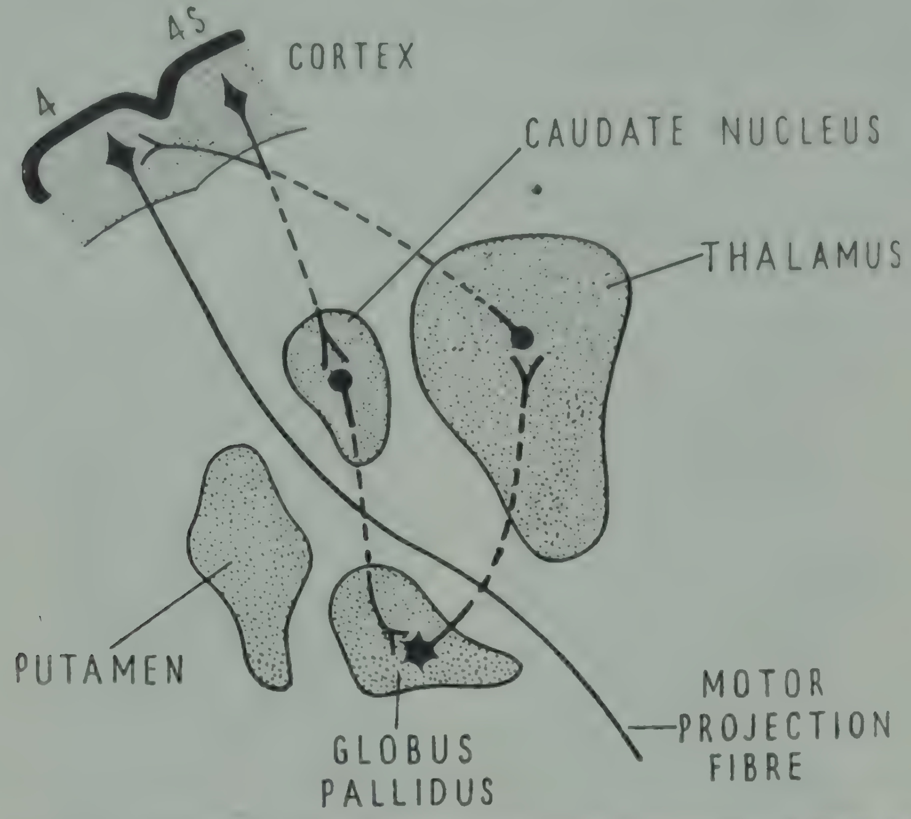

4S
4
CORTEX
CAUDATE NUCLEUS
THALAMUS
PUTAMEN
GLOBUS PALLIDUS
MOTOR PROJECTION FIBRE

FIG. 214. To illustrate path by which excitation of the suppressor strip 4s results in inhibition of the cortical motor area 4.

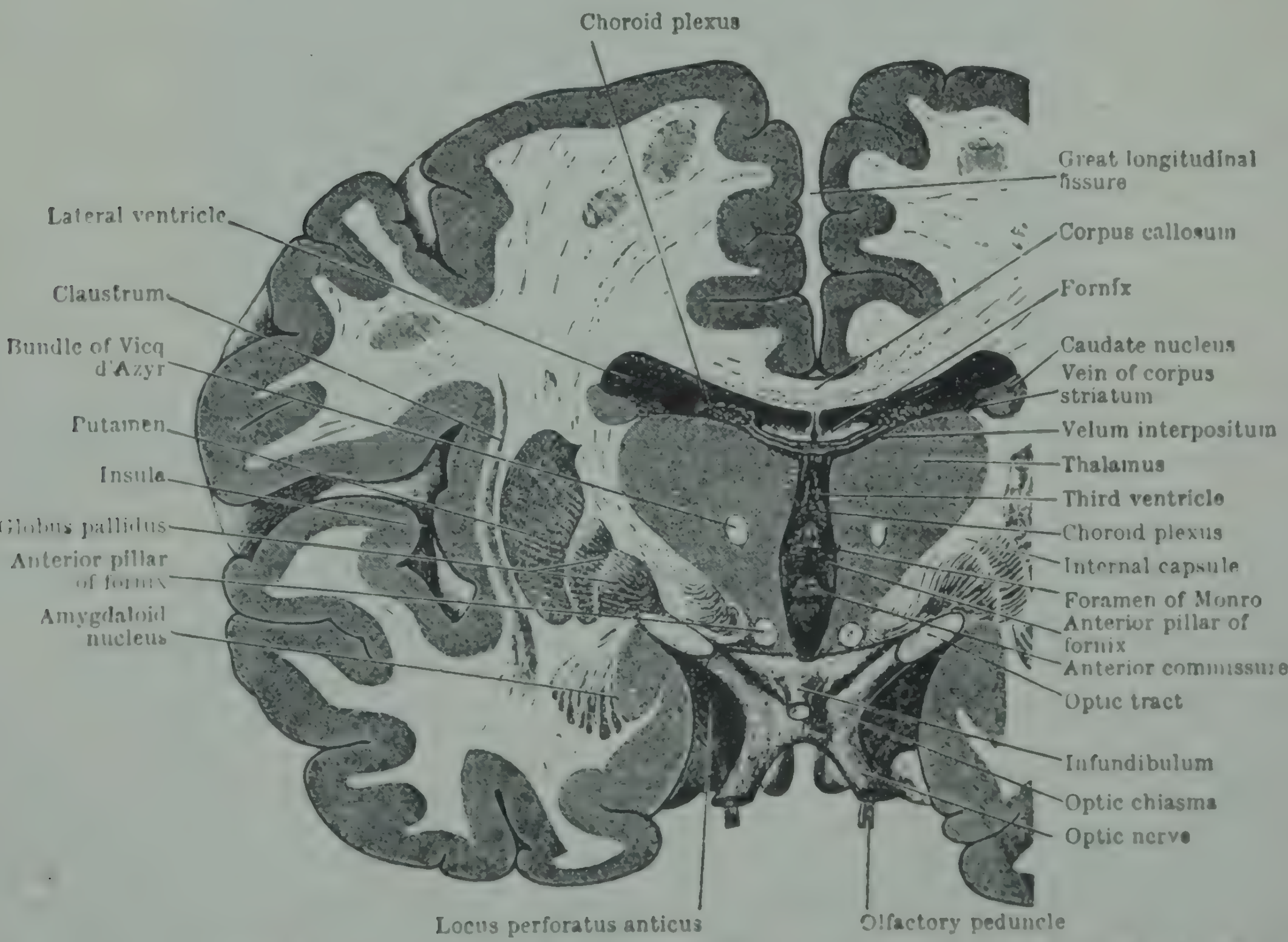

Choroid plexus
Lateral ventricle
Claustrum
Bundle of Vicq d'Azyr
Putamen
Insula
Globus pallidus
Anterior pillar of fornix
Amygdaloid nucleus
Great longitudinal fissure
Corpus callosum
Fornix
Caudate nucleus
Vein of corpus striatum
Velum interpositum
Thalamus
Third ventricle
Choroid plexus
Internal capsule
Foramen of Monro
Anterior pillar of fornix
Anterior commissure
Optic tract
Infundibulum
Optic chiasma
Optic nerve
Locus perforatus anticus
Olfactory peduncle

FIG. 215. Coronal Section through the Cerebrum so as to cut through the Lenticular Nucleus; Posterior Surface of the Section depicted. (CUNNINGHAM.)

---

* DUSSER DE BARENNE and MCCULLOCH. *J. Neurophysiol.*, 1941, **3**, 311.

but they no doubt have other important functions, and very flexible and varied connections. Electroneuronography is such a sensitive method that it tells us whether *any* neurones connect in a given way rather than whether a given path is a quantitatively important one, and so exposes to the risk of finding what we look for.

The motor projection areas 4 and 6 also have connections, *e.g.* with putamen and globus pallidus, and *viâ* these with the substantia nigra, nucleus of Luys, and hypothalamus, and other extra-pyramidal stations, but it is not yet known whether these connections mediate any controlling influence on the cortex in the way that the caudate nucleus does.

The balance of probability seems to be that muscular responses are controlled in a variety of ways—by reflected action on the cortex, by action through extra-pyramidal projections on the lower motor neurone, and, of course, through spinal reflex mechanisms, such as the stretch reflex.

Lesions of the Motor Paths in Man. The pyramidal path comprises three neurones. The first, called the *upper motor neurone*, starts in a pyramidal cell of the cortex and ends at the second, an intercalated neurone in the dorsal horn of the cord, which passes on to a motor neurone in the ventral horn, and this third, or *lower motor neurone*, comprises the spinal motor neurone and its axon enters one of the motor nerves. It has long been recognised by clinicians that the effects of section of the upper and lower motor neurones differ greatly. Lower motor neurone lesions, *e.g.* by cutting the motor nerve, or by destruction of the motor nerve cell bodies, as in the results of poliomyelitis, produce permanent flaccid paralysis and subsequent wasting of the muscles.

Lesions of the upper motor neurone as a rule produce quite other results. If the lesion is in the cortex the result will depend upon its position and extent. These have already been considered ; the main point is that at least partial recovery occurs, but spasticity appears and persists if the lesion is extensive.

Mere section of the pyramidal tracts apparently leads only to a flaccid paralysis,* whereas destruction of the motor path in the internal capsule by a hæmorrhage results in a spastic condition. The commonest site for lesions of the upper motor neurone in man is in the internal capsule, the cause being thrombosis in, or rupture of, the small arteries that supply it (apoplexy). The result, if on one side, is a paralysis of the muscles of the opposite side, the situation depending on the actual site and extent of the lesion. Often it is a general paralysis of one side (hemiplegia), the right-sided paralysis due to a left-sided lesion being accompanied by loss of speech. Respiration and other involuntary movements, *e.g.* some postural reflexes, are little affected, and muscles that act on both sides, *e.g.* eyes, trunk, etc., suffer only slight loss of function.

At first, in hemiplegia, the paralysed muscles are flaccid, but after a stage of shock has passed off, certain muscles become spastic. By spasticity is meant stiffness, highly active reflexes and a tendency to a clonic response on sudden stretching. The knee-jerk is more sustained than normal, a Babinski plantar reflex is obtained, there is often ankle-clonus, and characteristic limb attitudes, at first of extension in legs and flexion in arms, later of spastic flexion, also in legs. These are similar to the ones described under decerebrate rigidity (p. 274), and the explanation is no doubt similar, viz. that the rigidity is due to release from some higher control.

It is difficult to see why the upper motor neurone lesion should give a state of tonus so different from that resulting from lower motor neurone lesions, unless we admit that there is also damage to neurones other than those of the pyramidal tract but which also play upon the final common path of the lower motor neurones. This is generally admitted for the condition of decerebrate rigidity, and in totally decorticate animals, etc., where quite complex motor reflexes can be carried out. We must suppose, therefore, that the spasticity is the result, not of the interference with the pyramidal path, but of damage to the extra-pyramidal paths which run near to them, and which normally inhibit that overaction of lower reflex centres, which in its most developed form is recognised as decerebrate rigidity.

---

* Marshall. *Arch. Neurol. Psych.*, 1934, **32**, 778 ; Tower. *Brain*, 1935, **58**, 238; 1936, **59**, 408.

The Part played by Afferent Impressions in the Co-ordination of Muscular Movements. As a result of the carrying out of muscular movements, impulses are set up in the proprioceptive structures, including those of muscles. These secondary afferent impulses in their turn affect the excitability and the activity of the motor neurones, and are important in regulating the movements, whether these be aroused reflexly or volitionally.

When, in a mammal, a contraction of the quadriceps extensor is induced by a single shock applied to the intact ant. crural nerve, the curve obtained shows a relaxation time much slower and more prolonged than when the cut nerve is similarly excited, or when the posterior roots alone are divided. The contraction of the muscle, when its afferent path is intact, seems to develop reflexly in the muscle itself a condition of tone which damps down the relaxation. In the dog, after section of the afferent nerves of one hind limb, this limb is not at first used for walking ; it is kept more or less flexed at hip and knee, and later, when it is employed in walking, it is lifted too high with each step. After division of the afferent fibres of both limbs, these appear as if they were affected with motor paralysis. At first during walking, the fore limbs simply drag the hind limbs after them, though later, as the hind limbs are drawn along, they make alternate movements and may ultimately afford a certain amount of support to the body.

Still more striking effects are observed in complete deafferentation of the fore limb in monkey or man. The limb is never used in climbing, nor in the taking of food. That the peripheral motor mechanism is intact is shown by the fact that stimulation of the appropriate area of the cerebral cortex in such animals elicits at once a movement of the hand or limb. It seems, however, impossible for the cortex to *initiate* such movements in the absence of all afferent impulses arriving from the limb.

In these phenomena of sensory paralysis we are dealing with the effects produced by the deprivation of two distinct classes of afferent impressions, viz. those from the skin and those from the deep structures and muscles. The phenomena due to these two factors may be studied separately. Cutaneous insensibility of a limb of an animal causes very little disturbance of locomotion, the alternate movements of which seem to be regulated by the stimulation of deep structures and muscles. If, in the monkey, all the afferent brachial roots except the last cervical, which supplies cutaneous sensations to the whole hand, be divided, the monkey uses the arm and hand both in climbing and in taking food. A marked ataxy of the movement is, however, observed. Whereas the normal monkey, in taking grains of rice out of the observer's hand, exhibits perfect precision of movement, the monkey with only cutaneous sensibility remaining grasps clumsily with the whole hand, and the arm sways as it is put out, often missing the object aimed at altogether. These effects are due to the loss of sensations from the deep structures and muscles. Division of the afferent nerves from the muscles gives rise to a condition of tonelessness, and the passive mobility of the joints is greater than usual, so that the hip with the limb extended at the knee may be flexed to an abnormal extent.

The disturbance of co-ordination resulting from the cutting off of afferent muscular impressions is well seen in cases of tabes dorsalis (locomotor ataxy) in man, and to a slighter extent in cases of peripheral neuritis affecting chiefly the sensory nerves of muscles. The ataxic gait of such a patient is characteristic. There is no loss of power in the muscles, but there is loss of control. The patient is unaware of the position of his limbs and has to guide his walk by visual impressions : even then the movements are inco-ordinated. The contraction of every muscle is exaggerated, so that in walking the leg is first raised too high and then is brought down on to the ground with a stamp.

Such a patient may show "static ataxy," *i.e.* he is unable to stand with his feet together and his eyes shut. The afferent impressions from the skin of the feet, and from the muscles, appear therefore to be necessary also for the maintenance of static equilibrium. As the disease progresses the loss of control becomes more and more pronounced. The impressions are wanting which should cut short the contraction of a muscle when it has attained its optimum and which should inhibit the antagonists during the contraction and induce activity of the antagonists in successive alternation to those of the other muscles. In such a patient, therefore, walking finally becomes impossible, and, with well-nourished muscles and a motor path which is intact, he is condemned to pass the rest of his days in bed.

For Reference

Ass. Res. Nerv. Ment. Dis. (1932). Localisation of Function in the Cerebral Cortex. (Baltimore: Williams and Wilkins Co.)

Bills (1937). Fatigue in Mental Work. *Physiol. Rev.*, **17,** 436.

Brown, Graham (1927), in *Hdbch. d. norm. u. path. Physiol.*, **10,** 418.

Bucy (1944). The Precentral Motor Cortex. (Urbana. Ill.)

Hines, Marion (1929). On Cerebral Localisation. *Physiol. Rev.*, **9,** 462.

Hines, Marion (1943). Control of Movements by the Cerebral Cortex in Primates. *Biol. Rev.*, **18,** 1.

Lashley (1933). Integrative Functions of the Cerebral Cortex. *Physiol. Rev.*, **13,** 1.

Lorente de Nó (1943). Cerebral Cortex. Architecture, etc., pp. 274–301, in Fulton's Physiology of the Nervous System. (Oxford University Press.)

McCulloch (1944). The Functional Organization of the Cerebral Cortex. *Physiol. Rev.* **24,** 390.

Walshe (1948). Critical Studies in Neurology. (Edinburgh: Livingston.)

Walter (1938). The Technique and Application of Electro-encephalography. *J. Neurol. Psych.*, **1,** 359.

CHAPTER XXI

FUNCTIONS OF THE CEREBRAL CORTEX

From the evidence which has been reviewed it may be inferred that parts of the cerebral cortex act as final receiving stations for the afferent impressions from all parts of the body, including the organs of special sense, while other parts act as stations of origin for motor and other efferent impulses. A simple view of one of its functions would be that it acts as a centre in which, as in the spinal reflex centres, afferent impulses are reflected outwards as motor impulses. But we know also that the cerebral cortex is the seat of all the higher activities of the mind. In particular the frontal areas are concerned with these associative functions. The study of these belongs to the province of psychology, which employs methods different from those of physiology for its approach. Accompanying the passing-over of impulses from the afferent to the efferent side of the cerebral arc, it would seem that the state which we call consciousness is engendered. The states of consciousness glide continually from moment to moment in an unbroken stream of experience, consisting of a sharper focal content with a fringe of slighter definition, and leaving behind it a trace which we know as memory. By a process of attention we can single out parts of the stream of consciousness for closer focussing.

There seems but little doubt that our conscious experiences are the result of complex integrations of sensory impressions, which are assessed by being checked and compared with traces of previous experiences. Around each specific sensory receiving area is an associative zone, the function of which is believed to be concerned with the perception of the sensations aroused. Thus the visual sensory area is the area striata, Brodmann's area 17, and its stimulation leads to the production of sensations of light. Around it are areas 18 and 19, where these sensations are tested and integrated with tactile or other sensations so that out of the process visual perceptions are arrived at. Stimulation of these areas leads to hallucinations of definite objects, *i.e.* the sensation is intellectually classified and recognised.

There are various objective methods by which the physiology of cerebral activity may be studied, and they are of different significance. Among others, we have the study of reaction times, electroencephalograms, animal behaviour, conditioned reflexes, and the varied phenomena of disease.

Reaction Time. In studying reaction times we measure the time interval between the application of a stimulus and a voluntary response. The greater the complexity of the reaction, the longer will be the period which elapses between the stimulus and the response. Especially is this the case when the complex meshwork of neurones of the cerebral hemispheres is involved. The determination of the *reaction time* gives information as to the time relations of the psychical processes which are the representation in consciousness of the physiological changes occurring in the neurones of the central nervous system.

In most methods the application of the stimulus is arranged so as to close the circuit of a current which flows through an electro-magnet activating a lever which writes on a rapidly moving blackened surface. The reaction of the subject is arranged so that the movement activates a key by which the same current is opened. We thus obtain a tracing showing the moment of application of the stimulus and the moment at which the reaction takes place. If, *e.g.*, the sensory impression is to be from the skin, the current may be made to pass through the primary circuit of an induction coil, and wires be taken from the secondary coil to some part of the surface of the skin. In this case the signal may be started by opening the circuit, and the subject of the experiment will respond by closing the circuit by means of a spring key directly he feels the shock caused by the break of the primary circuit.

The average reaction times in man obtained with the different senses are as follows :—

Sight.	Hearing.	Electrical Stimulation of Skin.
0·19 to 0·22 sec.	0·12 to 0·18 sec.	0·12 to 0·20 sec.

The two figures given for each case are the extremes obtained in different series of observations.

The times vary according to the condition of the person who is the subject of the experiment. They are lengthened by fatigue ; they are shortened up to a certain point by continued practice. Within limits also they are shortened by increase of the strength of the stimulus.

Reaction time applies only to movements carried out in response to isolated stimuli, or to stimuli repeated at irregular intervals. When the stimuli are rhythmic the lost time applies only to the first one or two of the stimuli. The observer is conscious of the interval elapsing between the physical event and his reaction, and anticipates the later stimuli so that his reaction becomes synchronous with the stimulus. This synchronism of stimulus and reaction characterises all rhythmic movements, such as dancing or the playing of an orchestra.

Dilemma. When the subject is required to distinguish between alternative stimuli, the reaction time is considerably longer. Thus, the wires from the secondary coil may be connected by a switch to two pairs of electrodes which are applied, one to the right and one to the left half of the body. It is agreed beforehand that the subject shall react, say, only to stimulation of the right side. The switch is removed from the observation of the subject and the stimulus is applied irregularly to one side or to the other. It is found that the additional neural processes involved in determining whether the stimulus is on the right side, and therefore should be followed up as agreed, adds considerably (on an average 0·06 sec.) to the length of the reaction time. It is possible to complicate the dilemma to almost any extent. Thus the experiment may be so arranged that either a red or a white disc appears, and the subject has to react with the right hand to the red disc and with the left hand to the white disc. In such an experiment the reaction time was found to be 0·154 sec. longer than the simple reaction time. A still more complex process would be involved in the experiment in which a word was spoken, and the subject had to speak some other word which had some association with the word which formed the stimulus, *e.g.* horse—mammal ; paper—pen, etc. In such an experiment the reaction time was found to be as long as 0·7 to 0·8 sec.

Electrical Activity of the Cortex. When the cerebral cortex of an anæsthetised animal is led off by two electrodes placed on its surface and connected with a suitable amplifier and recording oscillograph, potential waves may be recorded.* These are roughly divisible into large slow waves, at 1–4 per sec., and superimposed smaller waves. These waves are the resultant of large numbers of short waves in individual neurones, the response of each unit probably varying in duration according to conditions, from 0·1–0·01 sec., but of uniform size. Increasing or reducing the anæsthesia may alter the magnitude of the waves : injury causes a rapid series of waves, the frequency of which gradually dies down from 100 to about 40 per sec., and which appear to be due to synchronous pulsation in large groups of neurones (Fig. 216).

The cause of the synchronisation, and of the wave-character of the cortical potentials is largely a matter for speculation. It has been suggested that chemical environmental conditions, or polarisation changes in the cells, may be the explanation. Thus, a layer of nerve cells, oriented with axons centrally directed, is normally polarised with the axonic ends positive : this could behave like a polarised nerve membrane and local depolarisation could spread over the surface as a wave.†

The arrival of afferent impressions at the cortex results in the starting off of trains of potential waves which radiate from the special sensory area concerned. Thus, if light falls on the eye, rhythmic discharges spread out from the area striata,* and so on. There is, as we have seen, a point-for-point representation of the receptor surface on the cortical area which receives the impulses, and it is also found that the time-distribution of the electrical changes in the cortex is related to those in the receptors. The world around us, and our own body sensations, are thus translated into a spatially and temporally distributed pattern of electrical potential changes, and it is these patterns which are woven into our state of consciousness, and which brings us all the information we can get about the universe around us. It is not certain whether these waves are initiated in the afferent terminal fibres or in internuncial cells affected by them. They show fluctuations according to the general background of cortical activity on which they impinge. The rhythmic potential waves of this background tend to be increased locally on arrival of an afferent stimulus and this effect resembles that of direct electrical stimulation of the cortex.†

---

* Adrian and Matthews. *J. Physiol.*, 1934, **81**, 440.

† Libet and Gerard. *J. Neurophysiol.*, 1941, **4**, 438.

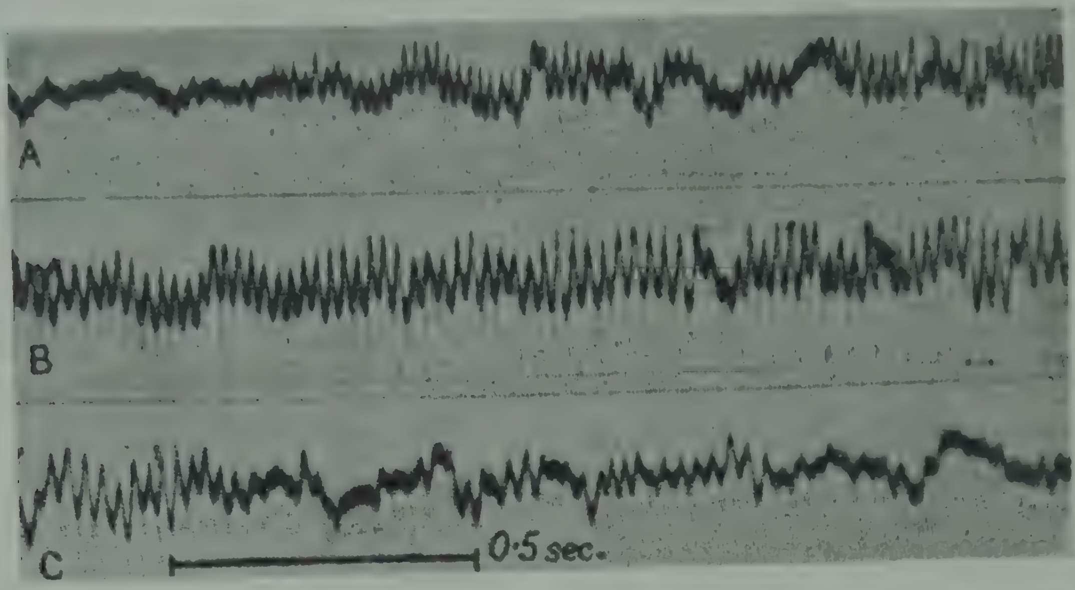


FIG. 216. Injury Discharge from Cortex of Anæsthetised Rabbit, punctured with fine Wire. A. Beginning, frequency = 64 per second ; B. Middle, frequency = 44 per second : C. End, frequency about 40 per second. (ADRIAN and MATTHEWS, *J. Physiol.*, 1934, **81**, 440.)

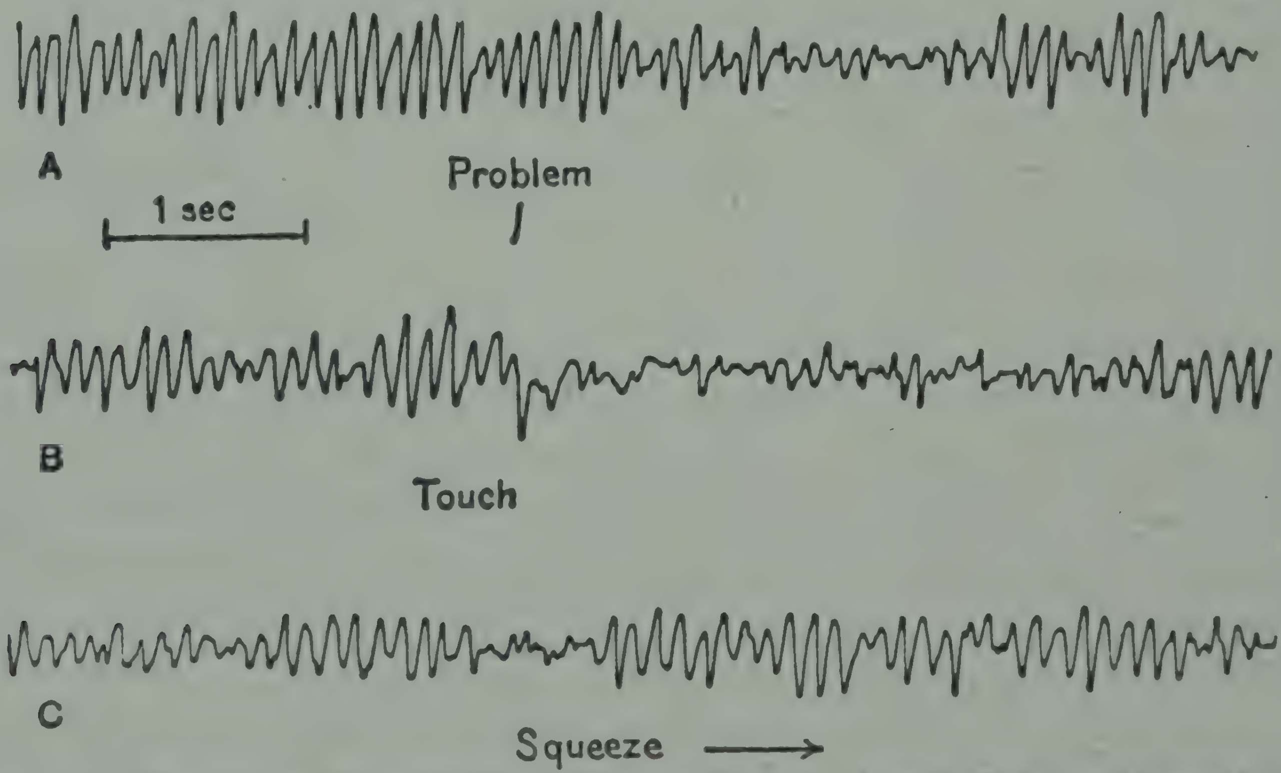


FIG. 217. "Berger" rhythm from Human Scalp. A and B show depression of waves by non-visual activities. At A, a mental arithmetic problem. At B, a touch on nose with cotton wool. C. Squeezing pliers tightly shows that muscular efforts do not affect rhythm greatly. (ADRIAN and MATTHEWS, *Brain*, 1934, **57**, 355.)

Electric stimulation of the cortex usually starts waves which radiate out

---

[*] CLAES. *Arch. internat. Physiol.*, 1939, **48**, 181.
[†] ADRIAN. *Proc. R. S.*, 1939, **126B**, 433.

over the cortex, the size of wave and the extent of spread increasing by facilitation as the stimuli are repeated. As the stimuli are repeated at the same spot, also, it is found that the point from which the waves begin tends to drift to other positions as other neurones, by facilitation, become more excitable. After further repeated stimuli, in fact, the whole cortex may be involved and may be driven as by a pacemaker, from the area of origin, and at its rhythm. If two centres compete in the emission of cortical waves, the regular rhythm may break down into small, rapid and irregular oscillations. Centres of inhibition may also be set up and the inhibitory influence may likewise irradiate widely over the cortex. This is especially well shown in certain areas, called suppressor bands (*e.g.* in areas 8s, 23s, 24s), stimulation of which abolishes spontaneous electrical activity over the entire cortex.

Severe injury or convulsant drugs cause the pulsations to spread widely over the cortex and in different directions, and large irregular waves are often seen.

Somewhat different potential waves (electroencephalograms) can be detected when electrodes are applied to the human scalp.* It is the practice to lead off from frontal, parietal, temporal and occipital regions on each side. The potentials are very minute (less than 50μV.) and need great amplification. The important feature is frequency, not amplitude of waves. The chief are the *alpha waves* (Fig. 217). These occur at about 10 per second; they usually arise in the visual area of the occipital lobe.† They are present when the eyes are shut or when, with uniform visual fields, there is no attempt to see detail, and are reduced or abolished whenever the central field of vision has any detail, or in any mental state, visual or otherwise, which demands attention. It is significant that the latent period for disappearance of the α waves on opening the eyes is about 0·4 second, *i.e.* it is longer than the motor reaction time. Adrian and Matthews regard these waves as due to spontaneous rhythmic activity arising simultaneously in groups of neurones in an area of visual cortex when this is not engaged in definite function.

Sometimes waves at about 25 per second also appear. These, the *beta waves*, possibly originate in the sensori-motor areas. They tend to be inhibited by tactile stimuli. In sleep there are sometimes intermittent "bursts" of waves at 14 to 15 per second, which are called "spindles," and in deep sleep a slow random rhythm, the *delta waves*, at about 6 a second.

It is probable that these various rhythms are not due to the effects of afferent impulses arriving at the brain, but represent a tendency on the part of groups of nerve cells to emit regular synchronous potential waves.

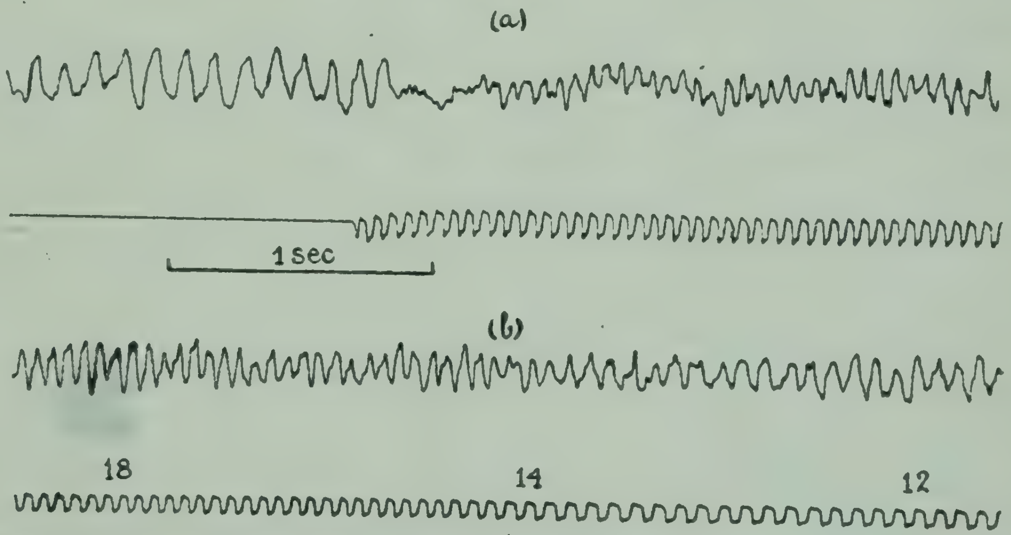


FIG. 218. Occipital Electroencephalograms from Man (*a*) Upper trace, EEG; lower trace, record of photo-electric cell, showing rate of flicker. Eyes were opened as flicker began. (*b*) Tracings as in (*a*): rate of flicker reduced from 18 to 12 per sec., and the EEG record follows it. (ADRIAN. *Lancet*, 1943, **245**, 33.)

This property is seen in almost any isolated portions of grey matter. Thus the isolated medullary centres of the goldfish show potential waves at a rate corresponding to that of the normal respiratory movements,* and portions of the frog's brain, especially the olfactory lobes, also show characteristic rhythms when isolated.†

---

* BERGER. *Arch. f. Psychiatr.*, 1929, **87**, 527; 1933, **102**, 454; KREEZER. *Amer. J. Psychol.*, 1938, **51**, 737.

† ADRIAN and MATTHEWS. *Brain*, 1934, **57**, 355.

In the brain *in situ*, moreover, it was found by Bremer ‡ that section across the mid-brain did not cause the cortical rhythms to disappear, although all afferents except olfactory and optic were removed thereby. These facts are of great importance inasmuch as they indicate that even in absence of stimuli from without, the cells of the cortex are not inactive, but tend over whole areas to fall into synchronous rhythms, the corresponding foci in the two hemispheres being also synchronous. There are some indications that mutual reinforcement by the cortex and the thalamus is largely responsible for maintaining the spontaneous rhythm.

If stimuli are given in a rhythmic manner, as by the exposure of the eyes to a flickering light, a community rhythm of the cells is very readily established, and corresponds with the rate of the flicker (Fig. 218).

ANIMAL BEHAVIOUR. Another objective method of investigation is by the study of the behaviour of animals as affected by conditions such as the extirpation of cortical areas. The methods mostly depend on the use of tests which involve the learning of something. Thus rats can learn to traverse a maze in response to a particular stimulus, *e.g.* light, or monkeys can acquire skill in unfastening latches, etc. After removal of the entire cerebral hemispheres nothing can be learnt, and even complex behaviour generally classed as instinctive, *e.g.* nesting of birds, may be lost. It is often supposed that when a habit has been acquired by voluntary effort, it ultimately becomes a reflex in which only the lower centres are concerned. The truth of the matter seems to be that cortical and sub-cortical processes react considerably one upon the other, and that many apparently cortical responses are really mediated by sub-cortical structures, though removal of the cortex does nevertheless modify or abolish such responses. Examination of animals in an ascending evolutionary scale shows that the cortex takes over to an increasing extent the functions of older structures, so that conclusions based on lower species are not necessarily applicable to higher ones.

One of the most striking results of animal behaviour experiments is the demonstration by Lashley that the effects of removal of cortical areas can be counteracted by training or retraining. Thus, rats can be trained to enter an illuminated alley and avoid a dark one. If, after training, the striate areas are removed the rat fails to pass the test and will never spontaneously recover the ability. Removal of up to 60 per cent. of other parts of the cortex leaves the ability unimpaired. Animals with striate areas destroyed can, however, be trained as quickly as normal rats to pass the test, and retraining of those who lost the ability by removal of the striate area also restores the ability. Such experiments, and similar ones on auditory centres, show that the idea of a simple path from receptor to cortex is quite inadequate.

Similar vicariation of function is shown in the motor paths of the cortex, and it has been found possible to train movements after removal of the motor centres primarily concerned with them, or after destruction of pyramidal tracts.

It may be concluded from experiments along these lines that there is less precision of localisation of functions in the cerebral cortex than the earlier work on localisation led us to expect, and that there seems to be a certain amount of mutual dependence between parts of the cortex. While there are certain fairly precise fields of local representation, or of special function there appear also to be wider and less precisely defined areas between which there is some latitude with respect to function. It is a surprising fact tha both frontal lobes can be removed in man without producing any very profound change (p. 301).

---

* ADRIAN and BUYTENDIJK. *J. Physiol.*, 1931, **71**, 121.

† GERARD and YOUNG. *Proc. R. S.*, 1937, **122**, 343.

‡ BREMER. *C. R. Soc. Biol.*, 1935, **118**, 1235, 1241.

For Reference

BRICKNER (1936). The Intellectual Functions of the Frontal Lobes. (Macmillan & Co.)
JACOBSEN (1939). The Effects of Extirpations on Higher Brain Processes. *Physiol. Rev.*, **19**, 303.
LASHLEY (1923). Integrative Functions of the Cerebral Cortex. *Physiol. Rev.*, **13**, 1.
WALTER (1938). The Technique and Application of Electro-encephalography. *J. Neurol. Psych.*, I, 359.

CONDITIONED REFLEXES

Pavlov introduced a method for investigating the functions of the higher centres, which is capable of very wide applications. It is applied to perfectly normal animals, and affords objective means of observation of the effects and interaction of sensory stimuli. Nevertheless, the conclusions drawn have not been universally accepted.

THE UNCONDITIONED AND THE CONDITIONED REFLEX. All reflex activities can be divided into two large groups: the *inborn*, and the *acquired* reflexes.

The inborn reflexes are characteristic of the central nervous system of any whole class of animals: they are transmitted by heredity, and their formation is therefore independent of the conditions under which any individual animal may live. Their strength may vary in the course of the animal's life: some of them do not appear until a definite age is reached, some disappear in senility, but in each case always to reappear in the next generation. To express the permanently stable character of these reflexes and their slight dependence upon surrounding conditions, Pavlov gave to the inborn reflexes the name of *unconditioned reflexes*. All spinal and bulbar reflexes are unconditioned reflexes: the knee-jerk, the stepping reflex, the secretion of saliva caused by administration of food, the flying reflex of birds, most of the postural, vascular and lower sexual reflexes, are included in the group of unconditioned reflexes. Some of these reflexes may be extremely complex and involve a series of successive reflexes, or be of an elaborate integrative nature, forming chains of reflexes following one another and presenting different groupings. What are generally known under the rather ill-defined name of "instincts" are probably nothing but a very complex integration of these inborn or unconditioned reflexes.

The acquired reflexes are peculiar to a single individual animal and may be entirely absent from the rest of the individuals of the same class. They are acquired in the course of the animal's existence, as a result of adaptation to changes in the surroundings. The organism at birth presents merely greater or lesser possibilities for the formation of these reflexes, but their actual formation depends only on the conditions under which the individual animal lives. They disappear more or less quickly when the conditions of the animal's life become so changed that it has no more use for the previously acquired reflexes; instead of the lost ones it forms new reflexes. Hence the name given to them of *conditioned reflexes*. However, it is important to remember that the conditions of life among individuals of the same class of animals may be identical in many respects; therefore it is not surprising that, besides possessing all the inborn unconditioned reflexes, these animals will also have many conditioned reflexes in common. The following example may serve as an illustration. Meat when placed in the mouth of any dog causes a reflex secretion of saliva and a reflex act of deglutition. It is common knowledge that salivation will also be produced if meat is only shown, or given to smell. The latter reflex will be found just as universally amongst the carnivorous animals as the former. Food therefore, besides its property of causing these reflexes when directly stimulating the taste, can cause a similar effect through the optic and the olfactory nerves. Only by means of special experiments was it possible to show that secretion caused by the application of the stimulus to the taste buds is based on an unconditioned reflex, while that excited by the *stimulation of the visual and olfactory* receptors is a conditioned, *i.e.* acquired, reflex, and is common to all *carnivorous animals only because the conditions of their life are identical in respect of their food.*

The Formation of Conditioned Reflexes. Conditioned reflexes never originate spontaneously. They only develop in association with another previously established reflex. In the simplest case, conditioned reflexes are based upon an inborn reflex and, since conditioned reflexes are not hereditary, we must look upon all conditioned reflexes as being an associative development of the inborn reflexes which ultimately lie at their root. Endless numbers of conditioned reflexes can be based upon one single unconditioned reflex. Conditioned reflexes of the second, third and higher orders can be developed in association with the primary conditioned reflexes, thus forming a complicated collection of nervous activities which are governed by education, experience and adaptation to the secondary conditions. The higher the organization of the animal's nervous system the greater the number and complexity of the conditioned activities it presents.

The experimental study of conditioned reflexes has dealt up to the present mostly with the simplest kind of reflexes which are developed in association directly with the unconditioned reflexes (conditioned reflexes of the first order). When a stimulus which is quite neutral, in that it does not by itself originate any kind of reflex activity of the animal, is repeatedly used during the execution of an inborn reflex (this latter of course in response to its proper stimulus), it loses its neutral character and acquires the property of bringing about by itself the same result as the inborn reflex. The hitherto neutral stimulus is transformed into a *conditioned stimulus*: it acquires the properties of the stimulus which initiated the reflex, and this can now be set into operation by the new accessory stimulus.

An example will illustrate the point. It is well known that food or acid when placed in an animal's mouth causes a reflex secretion of saliva, which in a dog with a salivary fistula can be collected and measured. This secretion is an inborn reflex, and does not require the presence of the cortex, since it can be observed after decerebration, or even on a bulbo-spinal preparation. Now if a second stimulus, say a sound (or any other stimulus), is applied to the animal simultaneously with feeding, and this combined stimulation is repeated several times, we find that the sound acquires stimulative properties and, when used alone without feeding, causes a secretion of saliva. A new reflex has thus been formed and a new nervous association established. That particular sound in respect of the salivary gland becomes a stimulus. Any other effector organ of the animal besides the salivary gland, or any sort of stimulus, can be used for the establishment of conditioned reflexes. On account of the comparative simplicity of measuring the reflexes the salivary glands have been most used as effectors. Conditioned reflexes have been established and studied also in the case of striated muscles, respiration, vasomotor system, pupil, deglutition, excretion of milk, contraction of the bladder, and intestinal and gastric secretion. It has been found possible to establish reflexes to visual, auditory, tactile, thermal, kinæsthetic, and olfactory stimuli.

Rate of Development of Conditioned Reflexes. Experiment shows that the development of conditioned reflexes is comparatively rapid. For example, a conditioned reflex was developed between the showing of food and the sounding of a note of 637·5 vibrations per second. The table on p. 325 shows that the reflex has been established by the thirty-first repetition.

Such a conditioned reflex as the above might be established within a few days, by repeating the combined stimuli of food and sound many times per day. When once such a conditioned reflex has been established, evidence of its persistence may be obtained for many years afterwards. If, however, the experiments be discontinued before it has been fully established the effects

Number of repetitions of combined stimuli.	Strength of the conditioned reflex when sound tried alone. Drops per 30 secs.	Latent period of the secretion in secs.
—	0	—
9	18	15
15	30	4
31	65	2
41	64	3
51	69	2

will be of a temporary nature. Suppose, for example, the above experiment had been discontinued on the ninth day, then, if tested on the fifteenth day, with no intermediate experiments, the probability would have been only a small secretion of saliva, indicating that the effect hitherto produced was fading away. If, on the other hand, the trials were continued until the fifty-first experiment, then a considerable effect might be obtained many months later. This forms the basis of "ingrained habits."

Laws of Conditioned Reflexes. Since the phenomena connected with conditioned reflexes are almost always investigated by means of some easily observed reflex act, such as salivation, the idea is apt to arise that conditioned reflexes only concern phenomena of this type. Such an idea is, of course, entirely erroneous. Such reflexes are chosen because by their use the observer has a simple and objective indication of what is happening in the animal's nervous system.

Summation. If, by means of separate applications, conditioned reflexes have been set up for two different kinds of stimuli, then, on subsequently applying both of them together, a greater effect may be produced than either would give alone. For example, suppose a conditioned reflex be set up by giving food and by sounding a whistle, and also in another series of experiments, carried on at the same time, by giving food and stimulating the back electrically, then, if subsequently the electrical stimulation is given and the whistle sounded together, the number of drops of saliva will be found to be the sum of those given by either stimulus alone.

Irradiation. Suppose that a conditioned reflex be established for a carefully-marked area of skin, and that by testing this area the strength of the conditioned reflex be measured. If now a slightly different area of skin be stimulated, it will probably be found that there is still a flow of saliva, though a smaller one than that obtainable if the correct spot had been stimulated. As spots farther and farther away from the correct one are tried, the number of drops decreases. A spreading of effect of this kind is called irradiation.

Specificity. Suppose that, having established a conditioned reflex for a particular gong, a gong of different pitch be sounded, then there may be either no flow of saliva or quite a considerable flow, according as the difference in pitch between the two gongs be perceptible or not. This phenomenon is called specificity.

Discrimination. Suppose that a series of gongs of different pitch are obtained and that two of them be experimented with in the following manner. The sounding of gong A is always accompanied by the giving of food, whereas the sounding of gong B is not. Then, on subsequently testing the animal for the existence of a conditioned reflex for these two gongs, it is possible to tell whether the animal has been able to distinguish between their pitches or not. Thus, if both of them are accompanied by secretion of the same number of drops of saliva on the average, then quite clearly they have sounded alike to the dog. By first testing with gongs which are very different in pitch, but in subsequent experiments using gongs which are closer and still more close, it is possible to ascertain precisely at what point discrimination ceases. Since very great care has to be taken to see that it is really nothing but the pitch which is the criterion that determines the dog's behaviour, tuning forks or organ pipes emitting pure tones will be preferable to gongs, and it would be necessary to be sure that no other quality changed at the same time, which would enable the animal to differentiate between them. Thus, one organ pipe might emit a slight high-pitched hiss when it was sounded. This might be quite inaudible to us but easily discerned by the dog. We should think that the dog was discriminating between the tones by their pitches, while in point of fact it was discriminating by the hiss in one of them.

REINFORCEMENT. Suppose that an animal in which a particular conditioned reflex has been established be not experimented with for several months, it will probably be found that the strength of the conditioned reflex has fallen somewhat from the value it had when the experiments ceased. This phenomenon is called decay. Experiment shows that a very few applications of the conditioning factors suffice to restore the reflex to its original value. This phenomenon is called reinforcement.

LINKING. Suppose that an animal with a conditioned reflex established between salivation and the sounding of a gong is experimented with for a few days by sounding a gong and a whistle at the same time. Then on subsequently testing the animal with the whistle it will be found that the conditioned reflex has a strength far greater than could be accounted for by the few times that the sound of the whistle and the feeding have been associated. What apparently has happened is that the conditioned reflex established for one sound has in great measure been transferred to the second. This phenomenon is called linking.

EXTINCTION. Suppose a conditioned reflex to have been established between the salivary reflex and the blowing of a whistle. Then, if for a number of times in succession the whistle be blown but no food given, it will be found that the strength of the conditioned reflex quickly dies away. This phenomenon is called extinction.

Delayed and Trace Reflexes. It is clear that the fundamental experiments with the modifications just described can be elaborated in a number of different directions. One of these consists in varying the time that is allowed to elapse between the application of the conditioned stimulus and the subsequent unconditioned stimulus. For example, the sounding of a note and the giving of food may occur simultaneously, or any number of minutes may be allowed to elapse between the sounding of the note and the giving of the food. If they occur together the conditioned reflex is called *simultaneous*. If a short time is allowed to elapse between the two, they are called *delayed*, and if the time is longer they are called *trace reflexes*. Now, as one would expect, delayed reflexes take longer to establish and are more liable to fade than simultaneous reflexes, and the same is still more true of trace reflexes.

Now it is clearly possible to vary in the opposite direction the time relations between the sounding of the gong and the giving of the food, namely, to give the food first and then to sound the gong. Experiment shows that unless the time be very short the establishment of a conditioned reflex becomes difficult or impossible. This is because the animal, as the result of repeated associations of the sounding of the gong and the giving of food, forms an association between the two, that is to say, the mechanism which sounds the gong is in some way connected with that which produces the food. If, on the contrary, the food is given first, so far as the animal is concerned subsequent events happening to the gong are of little or no interest.

Inhibition. EXTERNAL INHIBITION. Suppose that between the time of sounding the gong and the giving of food an additional unexpected stimulus be introduced, such, for example, as playing a gramophone record. Then it is found that this extra stimulus has a marked depressant effect on the strength of the conditioned reflex. The first application of this unexpected stimulus has a more powerful depressing effect than subsequent ones, as the following table shows :—

Strength of visual conditioned reflex . . . = 100 per cent.
Effect of gramophone, 1st application . . . = 10 ,,
,, ,, 2nd ,, . . . = 50 ,,
,, ,, 3rd ,, . . . = 65 ,,
,, ,, 4th ,, . . . = 85 ,,
,, ,, 5th ,, . . . = 90 ,,
,, ,, 6th ,, . . . = 94 ,,
,, ,, 7th ,, . . . = 100 ,,

This effect is known as external inhibition. Experiments of this nature make it clear that considerable pains have to be taken to eliminate the entry of unintentional external inhibition when initiating or testing the strength of conditioned reflexes. It has been found necessary to design special rooms (Fig. 219) provided with automatic means of giving the animal food at the required instant of time and for supplying the various forms of stimulation, gongs, whistles, etc., etc., so that adventitious stimuli of an unwanted kind are avoided so far as possible.

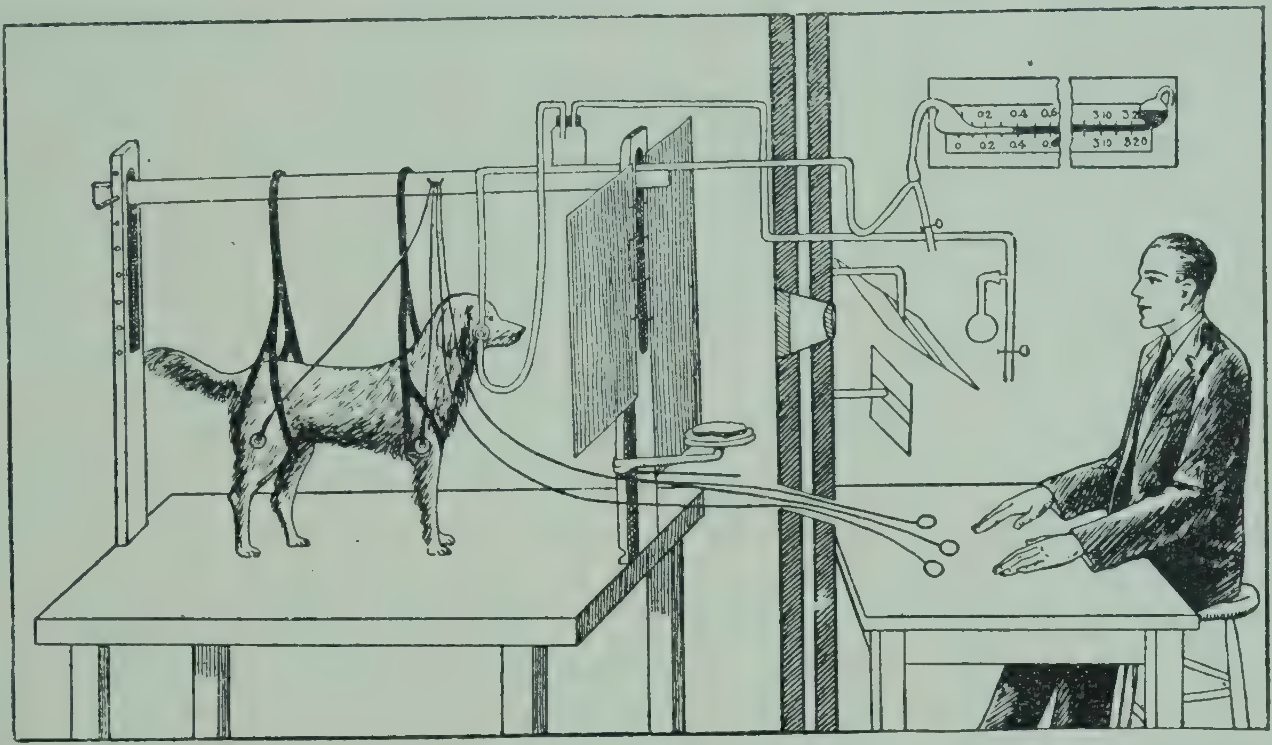


FIG. 219. Operator with various controls outside the cell observing animal inside the chamber by periscope. (PAVLOV.)

INTERNAL INHIBITION. Besides the type of inhibition described in the previous section, other varieties of inhibition are met with, some of which can be classed as internal inhibitions. For example, as the result of careful training a conditioned reflex may be set up for an organ pipe producing a note of 256 vibrations per second, whereas another similar pipe, producing a note of 244, is ineffective. In this case discrimination of the pitches of the two pipes has enabled the animal to react correctly to one and to inhibit the effects of the other. This is one type of internal inhibition.

EXPERIMENTAL NEUROSIS. Some of the dogs used in experiments involving rather fine and difficult discriminations showed signs of neurological stress, and ended in a neurotic state. They became highly excitable, bad-tempered, and generally unstable. Long treatment often restored them to apparent normality. It is of interest that it is just those mental processes which are concerned with discriminations which result in strain and fatigue in man.

EFFECT OF EXTERNAL STIMULI ON INTERNAL INHIBITION. We have already seen that adding an extra stimulus of an unexpected nature immediately before or after a stimulus to which a conditioned reflex has been set up may result in marked inhibition in the strength of the conditioned reflex. We must now consider the effects of putting an additional unexpected stimulus just in front of a stimulus to which internal inhibition has been established. For example, having caused the animal to discriminate between a tone of 256 to which he salivates, and another of 244 to which he does not, let us investigate what will happen if, just prior to sounding the pipe of 244, an unexpected additional stimulus is given to the animal. The result is that the animal salivates as if it was responding to the tone of 256 vibrations per second, and this clearly indicates that the internal inhibition may be temporarily removed by this unexpected extra stimulus.

SEAT OF INHIBITION. We must now consider whether the external inhibition which we have been considering develops primarily in the cortical area to which the normal stimulus goes, or whether it is associated more firmly with the cortical area to which the unexpected stimulus travels. Several observations seem to suggest that the inhibition arises within the cortical area to which the end organ concerned in the conditioned reflex connects. Suppose, for example, that as the result of repeatedly stimulating an area of skin and giving food, the conditioned reflex has been set up, and that a loud noise has been used to inhibit the conditioned reflex. Then experiment seems to show that the inhibitory after-effect will be much more marked in the cortical area representing the skin than in the cortical area representing hearing. The following experiment makes this clear. Suppose that a conditioned reflex has been set up for a particular area of skin, and that this has been inhibited by a loud sound so that the stimulation of skin only means to the animal to "expect food," whereas the loud sound and the stimulation of skin together mean "no food." Now, during a subsequent period another skin stimulus is used at the same spot for establishing a new reflex, for example, the contraction of a particular muscle instead of the secretion of saliva. Then, for the first time that the loud sound is applied at the same time as stimulation of the skin, there will be found to be inhibition of the movement of the muscle, just as there was inhibition of the flow of saliva. This clearly indicates that the association is between areas of cortex and not between the final paths used for the reflexes in question.

Conditioned Reflexes of Higher Orders. It is clear that by using conditioned reflexes in various combinations, sometimes using the same stimulus for two different reflexes, sometimes using a stimulus which has been used to inhibit one reflex to bring about the operation of a second one, we can link up conditioned reflexes in a very complicated way, only being limited by lack of time or of our own ingenuity. One example may be given. Suppose that a stimulus A be associated with the giving of food so that a conditioned reflex has been set up. Suppose that another stimulus B has been used to inhibit it, so that A and B together signify that no food is coming. Suppose now that for several times in succession the inhibitory stimulus B is associated with a third stimulus C which the animal has not previously experienced. Then on trying A and C together it will be found that C inhibits just as effectively as B did.

Removal of or Damage to Cortical Areas. Dogs which have survived the experimental removal of their cerebral cortex suffer a complete loss of the ability to form conditioned reflexes. Local damage only causes a temporary effect on conditioned reflexes if the damaged areas are of small size. Further, the damage appears to affect inhibitory stimuli more than excitatory ones. If, on the other hand, extirpation of the whole of the cortical area corresponding to a particular sense organ has been carried out, then many, but not all, of the effects previously produced by stimulating this sense organ are lost. Such conditioned responses as are retained are probably effected by sub-cortical mechanisms.*

Heredity of Conditioned Reflexes. One experiment to test this was performed as follows : Five puppies were removed from their mother and were fed entirely on milk. The first puppy had a conditioned reflex established between the giving of food and the sound of a metronome ; the second puppy to the stimulation of the skin by an induction coil; the third to the smell of camphor; the fourth to the appearance of a red light; and the fifth to a white light, which periodically went on and off. The conditioned reflexes developed quickly, and on testing were found to be quite stable. When the puppies were several months old they were shown a number of different objects, amongst them being pieces of meat and other foods which their ancestors had been acquainted with. No salivary activity was observed in the puppies. It did not matter whether the animals were merely shown the objects in question or were allowed also to smell them. The results were the same. Meat and rubber corks alike were indifferent to them. We see, then, that the salivary reflex to sight or smell, which might be supposed to be a hereditary characteristic possessed by all dogs, is in point of fact an acquired conditioned reflex.

---

* MARQUIS and HILGARD. *J. Comp. Psychol.*, 1936, **2**, 157.

Relation of Conditioned Reflex to Neurology. The main mass of neurological science has been built up on the basis of the neurone doctrine and its experimental interpretation by Sherrington. Conditioned reflex physiology has a large literature, and is clearly of great significance. But it has developed independently and has a special nomenclature, the relation of which to the tenets of general neurology is undefined. It is evident, for example, that the association between the conditioned and unconditioned stimuli on the one hand, and the effector path on the other hand implies a plasticity not so far revealed in the lower nerve centres. An attempt to harmonise these two fields of knowledge has been made by Konorski. He suggests that new synaptic connections can be established reversibly, and that these can be either excitatory in action, or, as in internal inhibition, inhibitory.

For Reference

H. S. LIDDELL (1943). Fulton's Physiology of the Nervous System, 491. (Oxford Univ Press.)

KONORSKI (1948). Conditioned Reflexes and Neuron Organisation. (Cambr. Univ. Press.)

PAVLOV (1927). Conditioned Reflexes. (Oxford University Press.)

PAVLOV (1928 and 1941). Lectures on Conditioned Reflexes. 2 vols. (London : Lawrence and Wishart Ltd.)

SLEEP

All higher animals have, during each day, one or more periods of rest in which there is a general reduction in the activity of the central nervous system, but especially of the cerebral cortex. The reduced, or delta, electric waves in the cortex are evidence of this. Adult human beings normally have one such sleep period in the twenty-four hours.

The need for, or onset of, sleep is announced by subjective feelings of drowsiness, dryness of the conjunctivæ, difficulty in keeping the eyes open, and general muscular relaxation.* When sleep sets in, the muscular relaxation increases, and before awakening, tonus returns. Many methods have been used for recording the movements made during sleep; the closing of electric contacts by movement of the bed springs, the use of smoked drum and lever, or cinema photography of the sleeper. Kleitman, Cooperman and Mullin † used a combination of some of the above methods with a machine which added together the vertical displacements of the centre of the bed springs. They also recorded the rectal temperature. Although most people make some movement every few minutes, little time is occupied by movements during sleeping, usually less than half a minute per hour. More time is spent in movement towards morning. There was no observable connection between rectal temperature and movement.

During sleep the knee-jerk disappears, but various superficial reflexes remain, and a positive Babinski sign is always obtained on stroking the sole of the foot. Urine secretion is reduced, and sometimes also gastric secretion and movement, but hunger contractions of the empty stomach may be increased. Vasomotor reflexes are conspicuous. The eyes are usually directed upwards, but any position may be taken up by either eye. The pupils are contracted, but contract still further to light. The heart rate is slowed, largely because of the general inactivity of the body, and the arterial blood pressure is lowered about 10 to 20 mm.* There is some dilution of the plasma and increase of blood volume.† Basal metabolic rate is reduced, but the respiration rate may be lowered, raised or unchanged.

---

* KLEITMAN. *Amer. J. Physiol.*, 1923, **66**, 67.

† KLEITMAN, COOPERMAN and MULLIN. *Amer. J. Physiol.*, 1933, **105**, 574.

The depth of sleep may be measured in a number of ways. One way is to take a record of the blood pressure, and to ascertain by experiment how loud a sound is necessary to cause a noticeable change in this. It is found that there are wide individual variations, but often sleep increases in depth until, two or three hours after the onset of sleep, it reaches its maximum. After this it decreases until about six hours after the onset of sleep, and then stays approximately uniform. Just before the time to wake up is reached, it becomes very light. Deep sleep is dreamless, and when dreaming does occur it is probably the result of some afferent impressions ; these are not subjected to the normal critical examination in the cortex, and so are usually grossly misinterpreted.

Prolonged sleeplessness in man leads to no very conspicuous objective changes, except as regards the results of certain psychological tests. Thus arithmetical problems are less accurately solved, and it is generally agreed that the driving of a motor car would be less reliable. In later stages it becomes difficult to stand erect, and hallucinations may occur. In order to keep awake after a certain time continuous muscular activity is essential. When very prolonged, sleeplessness leads to collapse and death. In such cases chromatolysis and shrinkage of cells in the frontal areas of the cortex are found.

As the result of experimental work on sleep, several hypotheses have been advanced to explain it. One of these, due to Pavlov, ascribes the phenomenon to a spread of internal inhibitory processes. If, for example, an animal has been trained to expect food half an hour after an auditory stimulus has been given (a trace reflex) and that this time interval be made longer and longer, the animal tends to go to sleep during the intervening period. When the time interval of the experiment is over the animal wakes up and exhibits the normal salivary reflex.

Kleitman regards sleep as a reversible inactivity of the highest cortical centres, which is due to a functional break between it and the lower parts. The highest levels are least organised and most sensible to fatigue. The break is due to a reduction in the number of afferent impressions, especially in the proprioceptive ones dependent on muscular tone. Hence muscular relaxation, whether voluntary or due to fatigue, is a necessary condition for it. This view receives important confirmation in the experiments of Bremer, who found that when nearly all afferent impressions were removed from the cerebral cortex in cats by section of the mid-brain, the head showed all signs, including the slow electric waves, of deep sleep. The diurnal alternation of wakefulness and sleep is probably a conditioned phenomenon.

Since the volume of the brain increases in sleep ‡ there can be no possibility that sleep is primarily due to cerebral anæmia, as often supposed. Some vascular phenomena may be related to it, however, since increase of pressure in the carotid sinus may under certain conditions induce sleep.§

Most neurologists now incline to the view that some special " centre " is concerned in the production of sleep, but opinions diverge as to the mode of its action. According to one school, led by v. Economo * the centre, near the tuber, could cause active inhibition of various higher centres, and this has been called a "sleep centre." Hess † was able to induce immediate sleep at any time in cats by stimulation of an area lying between the subthalamus and the thalamus, and near to the bundle of Vicq d'Azyr.

---

* LANDIS. *Amer. J. Physiol.*, 1925, **73**, 551.

† COOPERMAN. *Amer. J. Physiol.*, 1936, **116**, 531.

‡ SHEPARD. The Circulation and Sleep. New York, 1914.

§ KOCH *cit.* SCHWEITZER. Die Irradiation autonomer Reflexe. 1937 (Basel, Karger).

Others have thought it to be more appropriate to assume a "waking centre" the depression of which would cause sleep. Ranson ‡ indeed found evidence for such a "centre." He produced sleep, and sometimes catalepsy, in monkeys by bilateral lesions in the lateral hypothalamic area near the mammillary bodies, and similar results have been reported as a consequence of hypothalamic lesions in man. This is explained as due to the loss of the emotional drive which normally emanates from the hypothalamus, so that a stage of quiet and relaxation results. It is the converse of the state in the decorticate animal (p. 233) where the hypothalamic activity is uncontrolled and leads to extreme excitability and "sham rage." It was also found by Demole § that injections of 2 mg. of calcium chloride into the region of the tuber cinereum caused sleep lasting several hours. It could be reversed by injection of KCl. In pathological conditions of hypersomnia any lesions present are usually in the region lying between the corpora mammillaria and the third nerve nucleus.

Examination of the results of hypothalamic lesions in rats led Nauta to the conclusion || that there are in the hypothalamic region two neural mechanisms closely connected with sleeping and waking, and which may for convenience be called the sleeping and waking "centres." The "waking centre," in the caudal hypothalamic region, exercises a tonic influence on the cerebral cortex during the absence of external stimuli ; it is closely related to the sympathetic centres. This centre can be inhibited by a "sleep centre" situated in the preoptic region and related to parasympathetic centres. Sleep then occurs when this centre becomes active, provided that external stimuli at the time are not too powerful.

For Reference

KLEITMAN (1929). Sleep. *Physiol. Rev.*, **9**, 624.
RUSSELL BRAIN (1939). Sleep : Normal and Pathological. *B.M.J.*, (ii), 51.

---

[*] v. ECONOMO. *J. Neur. & Ment. Dis.*, 1930, **71**, 249.

[†] HESS. *Arch. f. Psych.*, 1936, **104**, 548 ; *Lancet*, 1932, **223**, 1259.

[‡] RANSON. *Arch. Neurol. & Psych.*, 1939, **41**, 1.

[§] DEMOLE. *Arch. f. exp. Path.*, 1926, **120**, 229.

[||] NAUTA. *J. Neurophysiol.*, 1946, **9**, 285.

CHAPTER XXII

THE CEREBELLUM

MORPHOLOGY

MANY attempts * have been made to systematize the anatomy of the cerebellum, with the result that the terminology is at present in a chaotic state. The cerebellum consists of a central part or vermis, and two lateral hemispheres. The vermis on the upper aspect is called the superior vermis and that on the inferior aspect the inferior vermis (nodule to tuber in old terminology). According to Larsell,† whose subdivisions we shall use, the cerebellum consists of two main divisions :—(1) a smaller *flocculo-nodular*

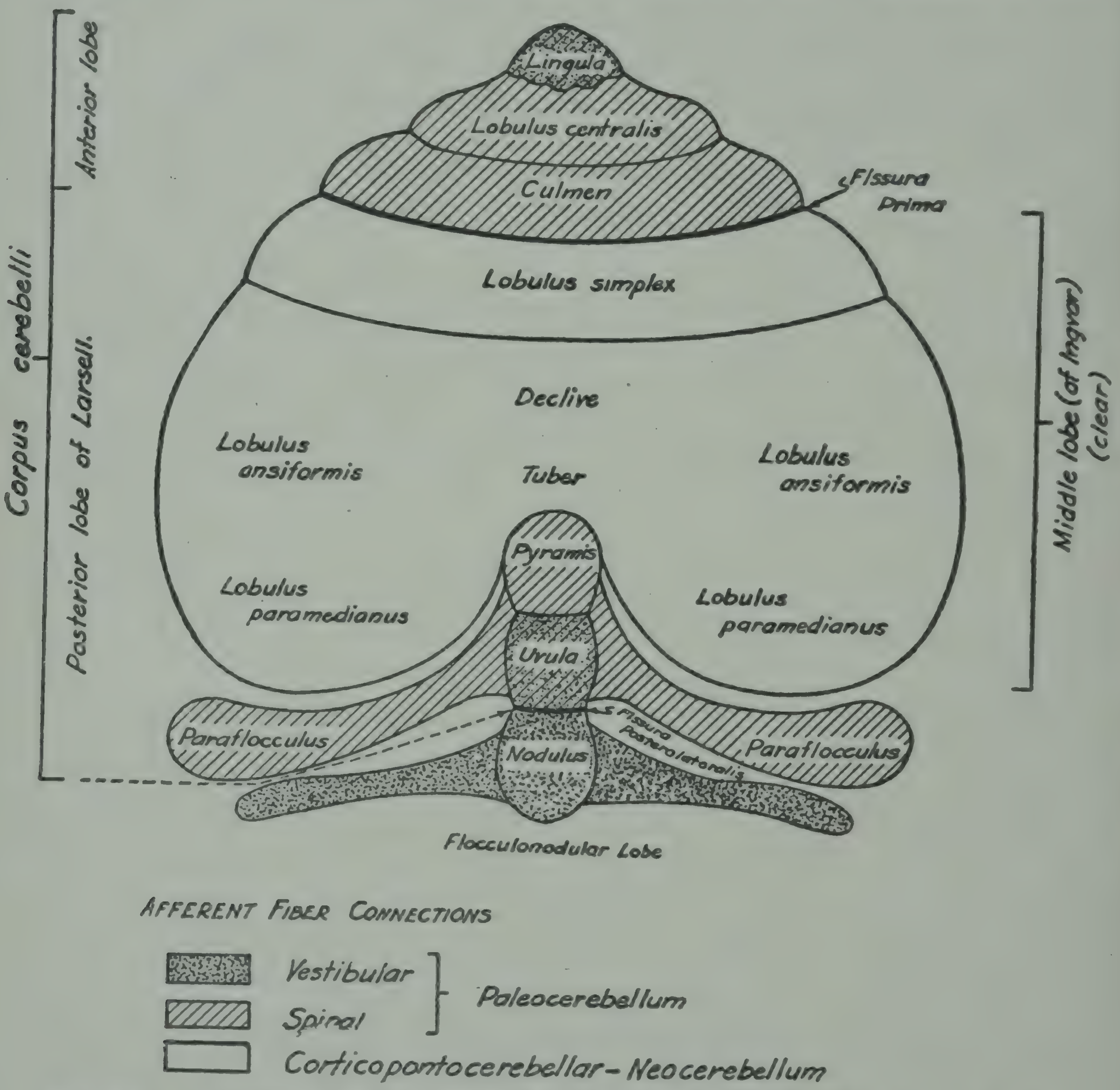

Anterior lobe
Corpus cerebelli
Posterior lobe of Larsell.
Lingula
Lobulus centralis
Culmen
Fissura Prima
Lobulus simplex
Middle lobe (of Ingvar) (clear)
Declive
Lobulus ansiformis
Tuber
Lobulus ansiformis
Pyramis
Lobulus paramedianus
Lobulus paramedianus
Uvula
Paraflocculus
Fissura Posterolateralis
Nodulus
Paraflocculus
Flocculonodular Lobe
AFFERENT FIBER CONNECTIONS
▨ Vestibular
▨ Spinal
Paleocerebellum
☐ Corticopontocerebellar - Neocerebellum

FIG. 220. Diagram of the Cerebellar Cortex of the Macaque, after Larsell. The surface is supposed to be opened out and spread in one plane. The principal divisions and afferent fibre connections are shown. The fissura posterolateralis separates the cerebellum into the small flocculo-nodular lobe and the larger corpus cerebelli. (FULTON. "Physiology of the Nervous System." Oxford University Press, 1938.)

---

* BOLK. "Das Zerebellum der Säugetiere" (1906) (Jena. Fischer). ELLIOT SMITH. *Anat. Anz.*, 1903, **23**, 368. INGVAR. *Folia Neurobiol.*, 1918, **9**, 205.

† LARSELL. *Arch. Neurol. Psych.*, 1937 **38**, 580. *J. Comp. Neurol.*, 1947, **87**, 85.

lobe, consisting of the flocculi, the nodulus and their peduncles, and (2) a larger *corpus cerebelli*, the two being separated by a fissure (not very conspicuous in man), the postero-lateral fissure.

The corpus cerebelli is again subdivided by the fissura prima, which runs transversely, into an *anterior lobe* and a *posterior lobe*, each further subdivided into smaller lobules by smaller fissures. Other folds run in such a way as to seem to divide the organ into a number of transverse laminæ. Ingvar recognises a middle and a posterior lobe in the corpus, but Larsell considers both as posterior lobe (Fig. 220).

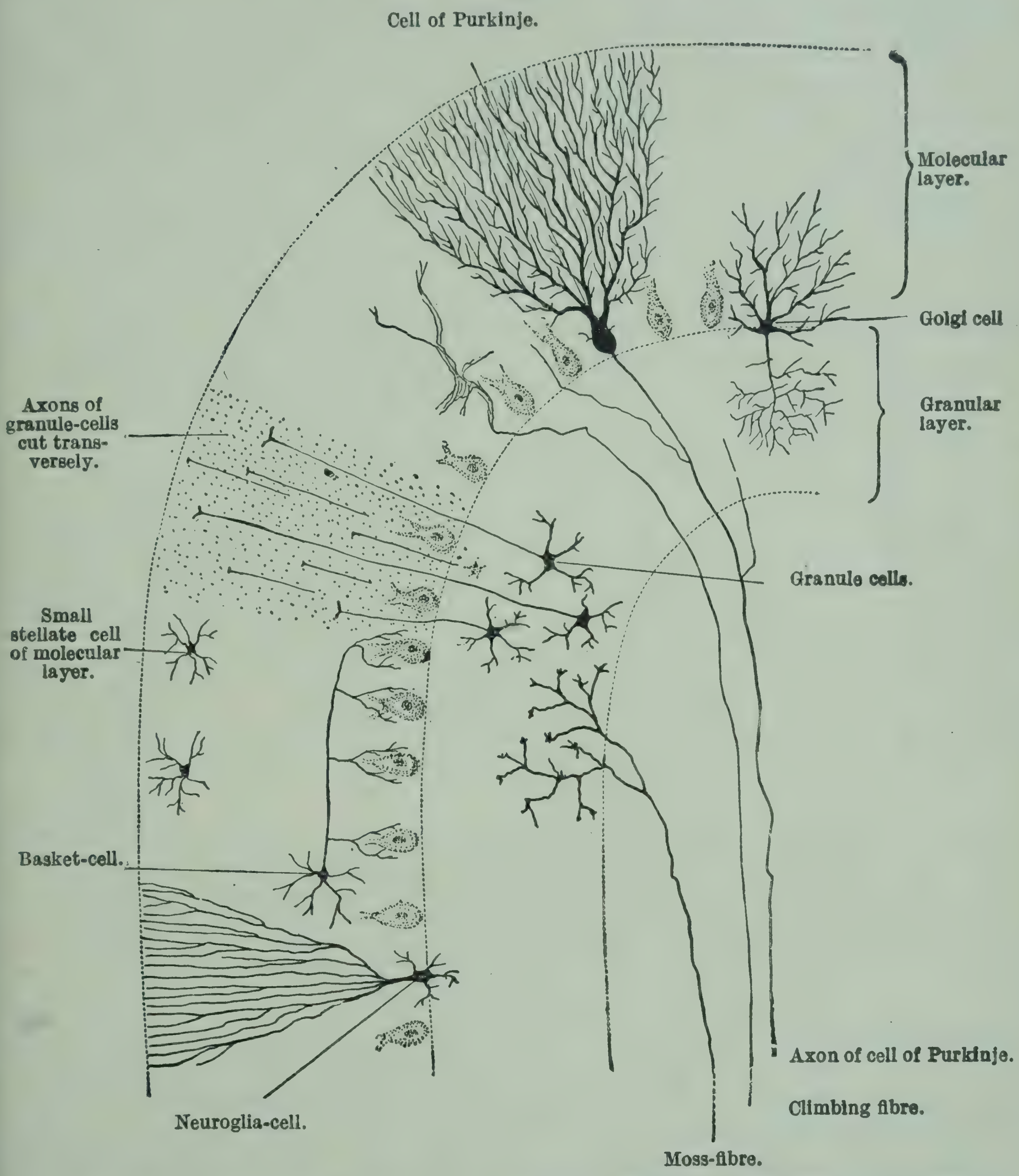

Cell of Purkinje.
Molecular layer.
Golgi cell
Granular layer.
Axons of granule-cells cut transversely.
Granule cells.
Small stellate cell of molecular layer.
Basket-cell.
Neuroglia-cell.
Axon of cell of Purkinje.
Climbing fibre.
Moss-fibre.

FIG. 221. Transverse Section of a Cerebellar Folium. (GRAY'S "Anatomy.")

The anterior lobe, the posterior part of the posterior lobe, and the flocculonodular lobe are the phylogenetically old parts, and hence may be called the *palæocerebellum*. They represent the original cerebellum as a structure developed chiefly in connection with the vestibular centres of the medulla. Their fibre connections as shown in Fig. 220 are vestibular and spinal. The anterior, and larger, part of the posterior lobe (= Ingvar's middle lobe) is greatly developed in higher animals and especially in primates: it is the *neo-cerebellum* and is chiefly connected, *viâ* the pons, with the efferent tracts from the cerebral cortex. The cerebellum connects with the rest of the central nervous system by means of three pairs of peduncles: the *superior peduncles* or brachia conjunctiva, the *middle peduncles*, which form the pons, and the *inferior peduncles* or restiform bodies. The cerebellum is covered with a grey cortex, within which is white matter.

The Cerebellar Nuclei are the three pairs of roof nuclei, viz. the *Nn. fastigii, globosus,* and *emboliformis* and the lateral *nucleus dentatus*, the last-named being the largest. Part of the dentate nucleus is connected with the palæocerebellum and part with the neocerebellum. All the nuclei are placed deeply and separated from the cerebellar cortex by the white matter. The general plan of connections is that the afferent fibres to the cerebellum go chiefly to its cortex: from this, fibres relay to the cerebellar nuclei, and from these again fibres pass out to other parts of the central nervous system, chiefly by the superior peduncles.

Histology of the Cerebellum. The cerebellar cortex is thrown into laminæ or folia by the transverse arrangement of the smaller fissures. Its substance shows three layers, and their structure is identical over the whole cortex. Outermost is the molecular layer, next a layer of Purkinje cells, and on the inside a granular layer. Beneath this is the white matter.

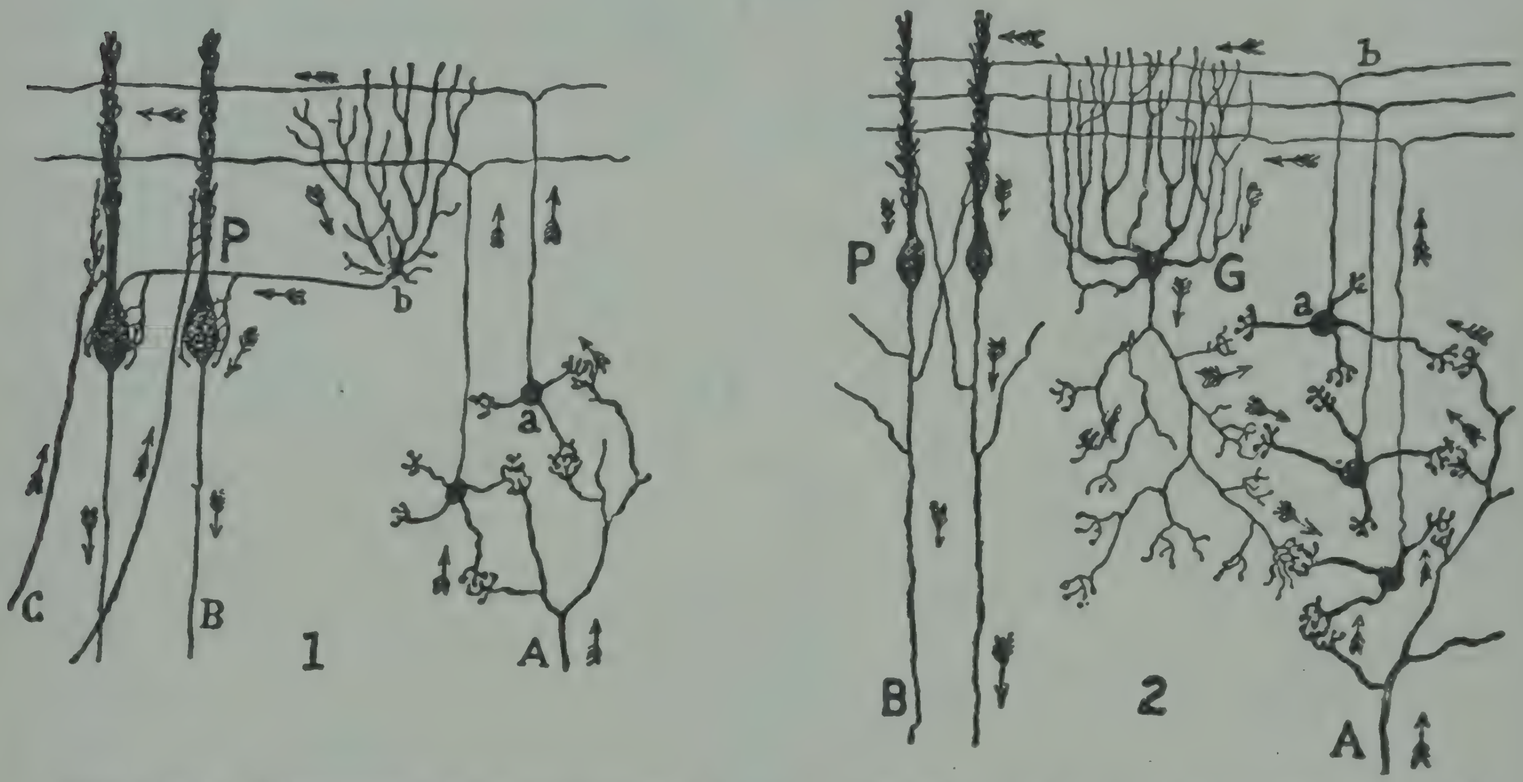


FIG. 222. Diagrams of Cajal showing two possible Paths for the Impulses carried by the Moss Fibres to the Cerebellar Cortex; the arrows indicate the direction of transmission of the impulses. 1. The unidirectional arc, and 2, the recurrent arc through the Golgi cells (G). A, moss fibre; B, axons of Purkinje cells; C, climbing fibres; a, granule cells; b, basket cells; P, Purkinje cells. (From LORENTE DE NÓ. *Arch. Neurol. Psych.*, 1933, **30**, 245.)

Starting at the deepest layer, the *granular layer*, this consists of small "granule" cells with short dendrons and long axons; the latter ascend to the molecular layer, where they bifurcate, the two branches then running parallel to the surface and parallel to the folia, *i.e.* transversely across the cerebellum (Fig. 221). Golgi type II cells are also found in the granular layer.

The *Purkinje cells*, of the next layer, have large bodies with freely branched dendrons that extend throughout the molecular layer and make synapses with the parallel axons of the granule cells (Fig. 221). The dendrons of the Purkinje cells are spread out in a plane which is right angles to that of the folia. Their axons pass to the cerebellar nuclei.

The *molecular layer*, in addition to the ascending axons of the granular cells, and their horizontal branches, contains small stellate cells and the so-called *basket cells*, the dendrites of which run at right angles to the folia, *i.e.* sagitally. They send out collaterals which end by basket-like arborisations round the bodies of Purkinje cells, so that several of these can be affected by one basket cell.

The *afferent fibres* entering the cerebellum go mainly straight to the cortex and are of two types :—(1) the *moss fibres* which end in the granule layer, and (2) *climbing or tendril fibres* which connect only with the Purkinje cells or their dendrons.

The moss fibres are derived chiefly from the spino- and olivo-cerebellar fibres and it will be seen that the arrangement of cerebellar neurones is such as to diffuse impulses from these widely over the cerebellum in a transverse direction by means of the parallel fibres, and also in a sagittal direction by the Purkinje cells and the basket cells by which these latter are associated.

The climbing fibres on the other hand, derived mainly from pontine and vestibular sources, carry impulses that are less widely diffused. From the climbing fibres they reach the Purkinje cells and the axons of these lead straight down to the cerebellar nuclei. Since some collaterals from the axons of these Purkinje cells are recurrent, there are no doubt opportunities for self-re-exciting circuits by which discharges are reinforced and prolonged. Similarly, the Golgi cells provide further opportunities for the reinforcement of impulses (Fig. 222). The histological evidence then, suggests that in the cerebellum incoming impulses are widely diffused and may also be augmented before their final emergence.

CONNECTIONS OF THE CEREBELLUM

The cerebellum receives the major part of the proprioceptive impulses from all sources, and acts as an important centre for their co-ordination. Its efferent fibres make connections with the motor system at all levels of the central nervous system. If any clue to its functions may be discerned in these connections, it is that the afferent connections of the palæocerebellum (roughly speaking, the vermis in older nomenclature) are vestibular and spinal, while those of the neocerebellum (the lateral lobes) are with the cortico-spinal and cortico-pontine projections.

The following are among the most important of its numerous connections (Fig. 223).

A. Afferent Connections

By the Inferior Peduncle

(1) *Vestibulo-Cerebellar Tracts.* From the vestibular nerve, and nuclei *viâ* the inferior peduncles to

(*a*) the flocculo-nodular lobe and part of the vermis of the anterior and posterior lobes (Fig. 220).

(*b*) the roof nuclei (chiefly fastigii).

(2) *Arcuate fibres* arising from cells of the nuclei gracilis and cuneatus, enter the inferior peduncle, and pass probably to the palæocerebellar cortex (not very important).

(3) *Olivo-cerebellar tract* fibres arise in the olive, cross over and pass up by the inferior peduncle to the cortex of the palæo- and neocerebellum.

(4) *Spino-cerebellar tracts.* (*a*) The *dorsal* spino-cerebellar tract fibres, with a contribution from the Vth cranial nerve, ascend by the inferior peduncle to reach the posterior parts (Culmen) of the anterior and the lobus simplex of the posterior lobe of the same side.

Exploration of the cerebellar cortex by micro-leads has shown that when the limbs are stimulated, or passively moved (most effective being pressure on the pads of the feet, with dorsiflexion of digits, as in standing), nervous impulses arrive at the mesial regions of the anterior lobe of the cerebellum *, impulses

---

* Adrian. *B. M. J.*, July 29, 1944, 137.

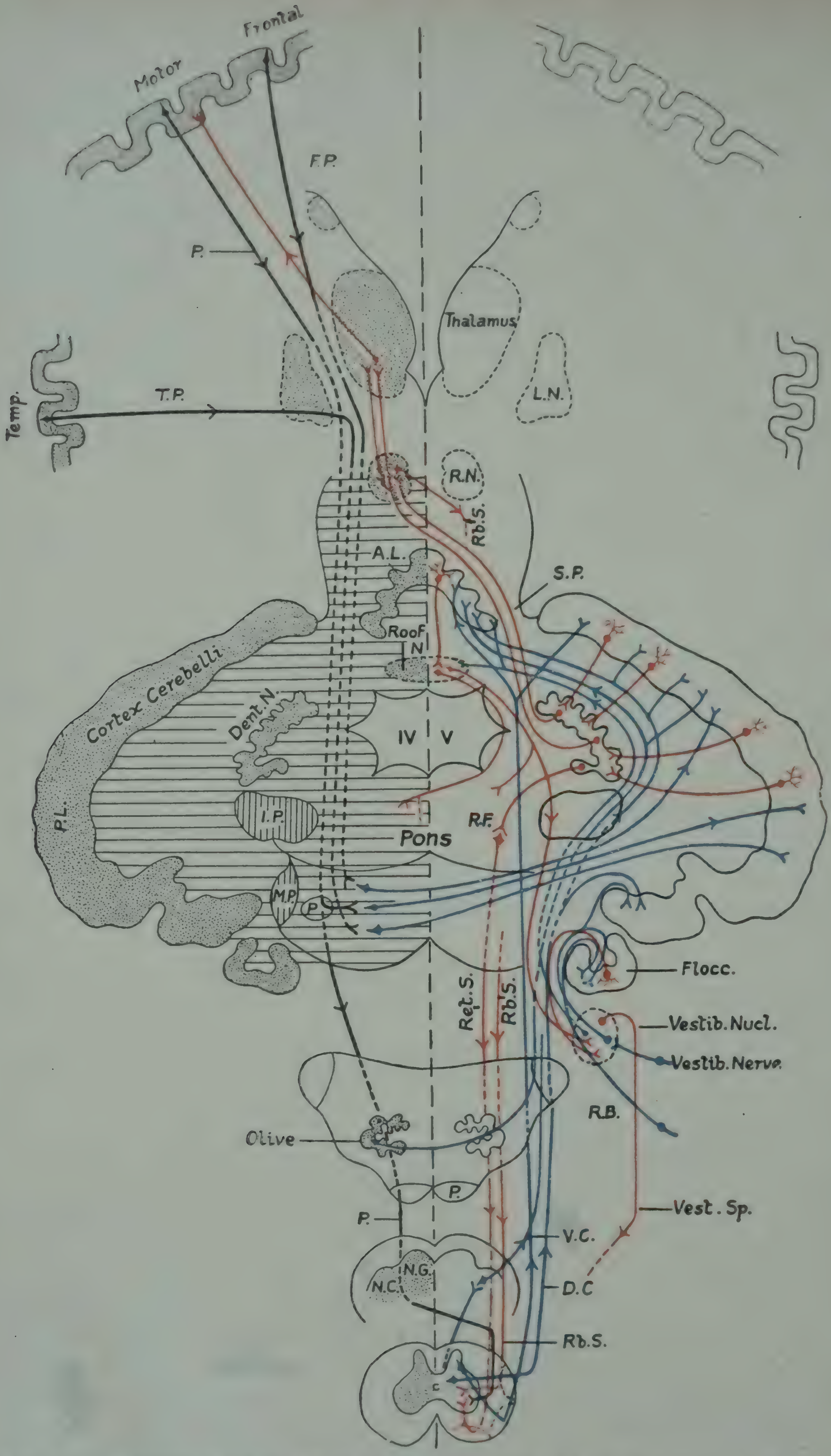

Frontal
Motor
F.P.
P.
Thalamus
Temp.
T.P.
L.N.
R.N.
Rb.S.
A.L.
S.P.
Roof
N
Cortex Cerebelli
Dent. N.
IV
V
P.L.
I.P.
R.F.
Pons
M.P.
P.
Flocc.
Ret. S.
Rb.S.
Vestib. Nucl.
Vestib. Nerve
R.B.
Olive
P.
P.
Vest. Sp.
V.C.
N.G.
N.C.
D.C
Rb.S.
c

FIG. 223. Diagram showing the principal connections of the Cerebellum. Afferent fibres blue, efferent fibres red. Cerebral efferent fibres black.

A.L., anterior lobe; D.C., dorsal cerebellar; Dent. N., dentate nucleus; Flocc, flocculo-nodular lobe; F.P., fronto-pontine fibres; Frontal, frontal cerebral cortex; I.P., inferior peduncle; L.N., lenticular nucleus; Motor, motor cortex; M.P., middle peduncle; N.C., N, G, nuclei and cuneatus gracilis; P, pyramidal fibres; P.L., posterior lobe; R.B., restiform body (blue fibres); Rb.S., rubrospinal fibres; Ret. Sp., reticulo-spinal fibres; R.F., formatio reticularis; R.N., red nucleus; Roof. N., roof nuclei; S.P., superior peduncle; Temp., temporal cortex; T.P., temporo-pontine fibres; V.C., ventral cerebellar; Vestib. Nerve and Nucl., vestibular nerve and nucleus from the hind limbs reaching the oral part, those from the forelimb the culmen, and those from the vibrissæ and face the lobus simplex (Fig. 224). There is thus, as in the cerebral cortex, a rough local representation of the proprioceptive regions projected on the cerebellar cortex, in order leg to face from "above downwards." The areas involved are larger when the proprioceptive regions to which they correspond are more important. Thus, the size of the culmen, which receives from the front limbs, grows progressively greater as we pass from pig to Shetland pony, and finally to man.

BY THE SUPERIOR PEDUNCLE

(*b*) The *ventral* spino-cerebellar tract passes up to the superior peduncle, and then turns back to enter the anterior lobe and roof nuclei of the same side.

BY THE MIDDLE PEDUNCLE

(5) *Cortico-cerebellar tracts.* Fibres from the cortico-pontine projections (= fronto-pontine and temporo-pontine) and branches from the cortico-spinal (pyramidal) tracts relay in cells of the *nuclei pontis*, axons from which conduct to the cortex of the neocerebellum of the opposite side. This is shown by the results of electrical exploration of the cerebellar cortex, while the cerebral motor cortex of the opposite side is stimulated.* Impulses are then observed to arrive at situations which show an orderly marshalling of local representation; the impulses reach more lateral regions of the anterior lobe than those conveyed from the spino-cerebellar tracts (Fig. 225). The greater part of the cortex of the posterior lobe seems to receive but few impulses from this source, or from the spino-cerebellar tracts.

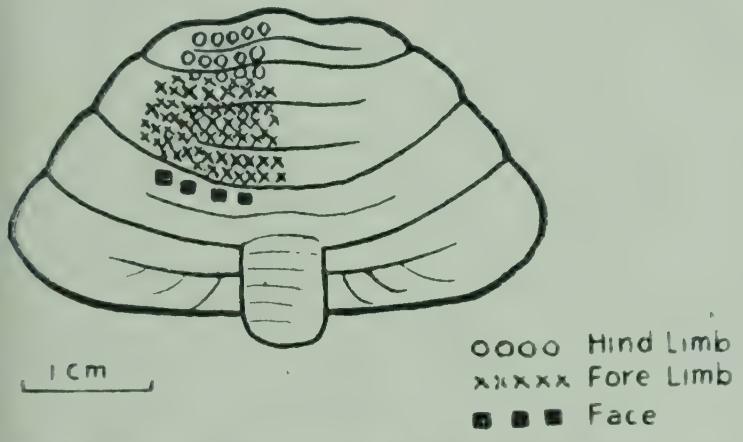


FIG. 224. Receiving areas for Spino-cerebellar Impulses in the Cerebellar Cortex of the Monkey. (ADRIAN, *B. M. J.*, 1944.)

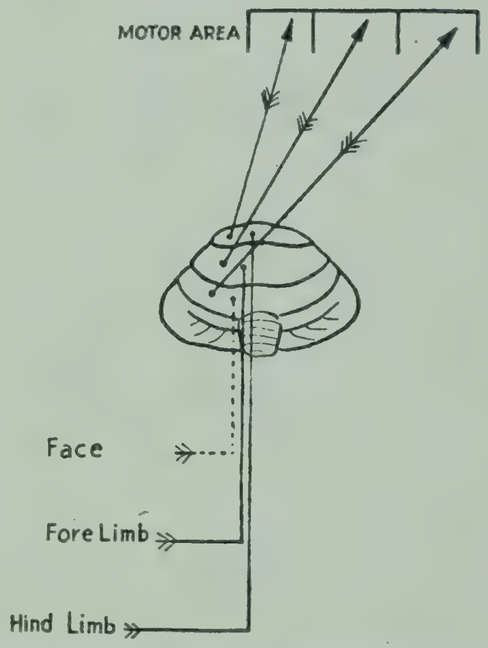


FIG. 225. Spino-cerebellar and Cortico-ponto-cerebellar Projections to receiving areas in Anterior Lobe and Lobus Simplex of the Cerebellar Cortex. (ADRIAN, *B. M. J.*, 1944.)

B. Efferent Connections. Fibres leaving the cerebellar cortex do so in two stages :—

I. FROM CEREBELLAR CORTEX to DEEP NUCLEI.

II. FROM DEEP NUCLEI to EXTRA-CEREBELLAR STRUCTURES. The chief paths outwards fall into three groups :—

(1) *Cerebello-tegmental tracts* which lead from the deep nuclei to tegmental motor nuclei of the reticular formation in the brain stem :—

(*a*) Fibres from roof nuclei *viâ* inferior peduncle to vestibular nuclei, posterior longitudinal bundle, reticular formation, reticulo-spinal tract, etc.

---

* HAMPSON. *J. Neurophysiol.*, 1949, **12**, 37.

(*b*) Russell's tract from roof nuclei to vestibular nuclei, posterior longitudinal bundle, etc., of the opposite side (Fig. 226).

(*c*) From roof nuclei and dentate nucleus (palæo-cerebellar part) by *superior peduncles*, fibres cross to the opposite side, and after sending collaterals to nuclei in medulla, pons and mid-brain, end in the large-celled part of the red nucleus. From these the rubro-spinal tract originates and this then crosses back again to the same side.

(2) *Cerebello-rubro-thalamic and cerebello-frontal tracts.* From the dentate nucleus (neocerebellar part) *viâ* the superior peduncles to the opposite side, and on to the small-celled red nucleus; relays from here, together with direct fibres from cerebellar peduncles, pass on to relay in the thalamus, or pass directly to the frontal and temporal cerebral cortex.

The efferent tracts thus fall, like the afferent ones, into two systems, one putting the palæocerebellum into relation with the motor and correlation centres of the brain stem and spinal cord, and one putting the neocerebellum into relation with the cerebral cortex.

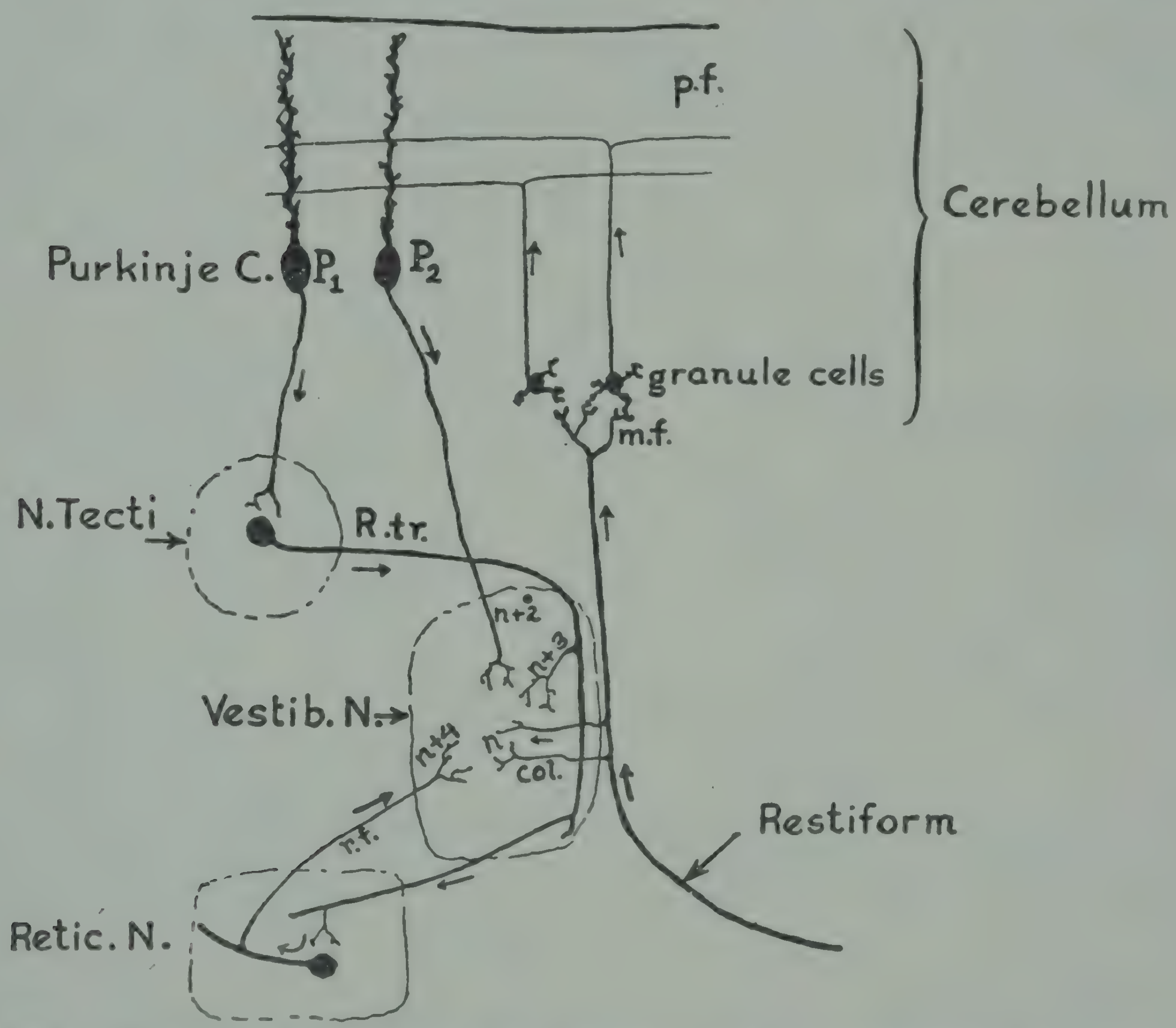


FIG. 226. Diagram of Pathways that Transmit the Proprioceptive Impulses carried by the Restiform Body to the Vestibular Nuclei. p.f., parallel fibres; m.f., moss fibre; R.tr. = Russell's tract, which gives off collaterals to the vestibular nuclei and ends in the reticular substance. From this, recurrent fibres pass back to the vestibular nuclei. (LORENTE DE NÓ. *Arch. Neurol. Psych.*, 1933, **30**, 245.)

It is important to note that the cerebellum is connected with the spinal cord of the same side and with the cerebral cortex of the opposite side.

FUNCTIONS OF THE CEREBELLUM

Knowledge of the functions of the cerebellum is derived from experiments on extirpation, on electro-neuronography, and on stimulation, and from clinical observations of the effects of disease or injury.

Extirpation. The effects of removal of the cerebellum differ according to the species of animal used and the extent of the ablation practised. It is necessary to distinguish between the effects of removal of the palæocerebellum and of the neocerebellum. The former affects equilibration and muscular tonus, the latter leads to impairment of voluntary movements.

In birds the whole cerebellum is a palæocerebellum, and when it is removed there follows a state of hypertonus of the extensor muscles.* Destruction of the anterior lobe alone (which has spinal connections) merely causes this extensor rigidity, but if the posterior lobe (with vestibular connections) is removed there is also defective equilibration, the head being bent backwards.

In mammals, removal of the whole cerebellum leaves sensation unaffected, modifies, though it does not prevent, voluntary movements, and causes disturbance of posture and equilibrium. There are great differences between species ; in lower mammals the effects are transitory, but in apes and men they may be long-lasting. Luciani † described three stages following cerebellar ablation in dogs :—(1) functional exaltation, (2) maximal cerebellar deficiency ; (3) compensation. In all stages there is inco-ordination of movement, or *ataxia*, and *tremor*.

In the first stage there is restlessness, with spasms of exaggerated tonus of the extensor muscles very similar to those of decerebrate rigidity. The positive supporting reaction of Magnus is greatly enhanced.‡ The results are probably due to release from cerebellar inhibitory influences ; an increase of rigidity can in fact be demonstrated on removal of the cerebellum from a decerebrate animal.

In the second stage, three symptoms appeared, called *asthenia* (weakness), *hypotonia*, and *astasia*—staggering gait and tremor on voluntary movement (intention tremor).

In the third stage, these deficiency symptoms are in part remedied by the improvement of voluntary control. The animal walks with the legs wide apart and so obtains a broad and secure base.

Removal of half the cerebellum, on one side only, leads to the appearance of the characteristic symptoms on the *same* side of the body, but recovery is much more complete than after total ablation.

Equilibration is not profoundly affected by removal of the cerebellum, and the labyrinthine postural reflexes are unaffected by it if the vestibular nuclei are not damaged.

Destruction of different parts of the cerebellum leads to results which are in accordance with what we know of their connections. Thus, if the flocculonodular lobe is destroyed there are disturbances of equilibrium, but not of voluntary movements § ; destruction of the anterior lobe causes great increase of extensor tonus as in the pigeon ; partial excisions of the anterior lobe show effects on face, fore-limb and hind-limb in order from behind forwards (as in Fig. 224) and most anteriorly, at the lingula, vestibular effects. Ablation of the neocerebellum results in grave disturbances of voluntary movements with hypotonia and, if the dentate nucleus is involved, with tremor and errors in range of voluntary movement.

Stimulation. No precise muscular movements are obtainable by stimulation of the cerebellar cortex.

Stimulation of the neocerebellum does, however, cause change in the motor areas of the cerebral cortex, as is shown by increase in its spontaneous electric rhythm || and by an increase in its excitability.¶

Stimulation of the anterior lobe does produce effects on the musculature, not contraction, but a relaxation of tonus, *e.g.* of that of decerebrate rigidity.**

---

* LANGE. *Pflüger's Arch.*, 1891, **50**, 615. TEN CATE. *Arch. Néerl. Physiol.*, 1926, **11**, 1 ; BREMER and LEY. *Arch. intern. Physiol.*, 1927, **28**, 58.

† LUCIANI. Il cervelletto, 1891 (Monnier, Florence). "Human Physiology," III., 1915, p. 419 (Macmillan, London).

‡ RADEMAKER. *Rev. neurologique*, 1930, **1**, 337.

§ DOW. *Arch. Neurol. Psych.*, 1938, **40**, 500.

|| EARL WALKER. *J. Neurophysiol.*, 1938, **1**, 16.

¶ BREMER. Le cervelet, Roger and Binet's *Traité*. Vol. **10**, p. 97.

** SHERRINGTON. *J. Physiol.*, 1898, **22**, 319.

The relaxation is usually of the antigravity muscles of both limbs of the same side, with no contraction of the antagonists, and also a smaller effect, or more rarely an augmentation of extensor tone, on the opposite side. If a crossed extensor spinal reflex is set up beforehand it is inhibited. The relaxation is followed, after stimulation, by a rebound contraction.* Sometimes extensor movements are obtained, but there is no exact localisation of these.

Stimulation of the flocculus causes movements and nystagmus of the eyes, which agrees with its known vestibular connection.

Stimulation of the deep nuclei † leads to various movements which are subjectively interpreted as being changes of posture.

The effects of stimulation are abolished by section of the superior peduncles.

Cerebellar Disease and Injury. The effects of irritative lesions of the cerebellum are usually too much complicated by accompanying effects to enable definite conclusions to be drawn.

In destructive lesions of the cerebellum there seem to be two clear-cut conditions, attended with different symptoms, namely, destruction of the palæocerebellar parts, and destruction of the neocerebellum. In neither is sensation affected.

Palæocerebellar Destruction. The palæocerebellum is insignificant in size in man, and hence is rarely affected alone. A midline tumour (medullo-blastoma) in children sometimes damages the flocculo-nodular lobe, and causes unsteadiness of gait, with walking on a wide base and a tendency to fall backwards. Tremor and hypotonia are not present, and voluntary movements of arms and legs are normal, though in walking there is inability to balance the body. A somewhat similar syndrome in adult men is seen in rare cases of primary palæocerebellar atrophy due to degeneration of the Purkinje cells, chiefly in the anterior lobe. Here again, the chief defect is in balancing the body on the legs; though voluntary movement of the legs is almost, and that of arms quite, unaffected. As in animals after removal of the anterior lobes, the proprioceptive reflexes concerned with equilibration are exaggerated, and the loss of equilibrium is no doubt due to this.

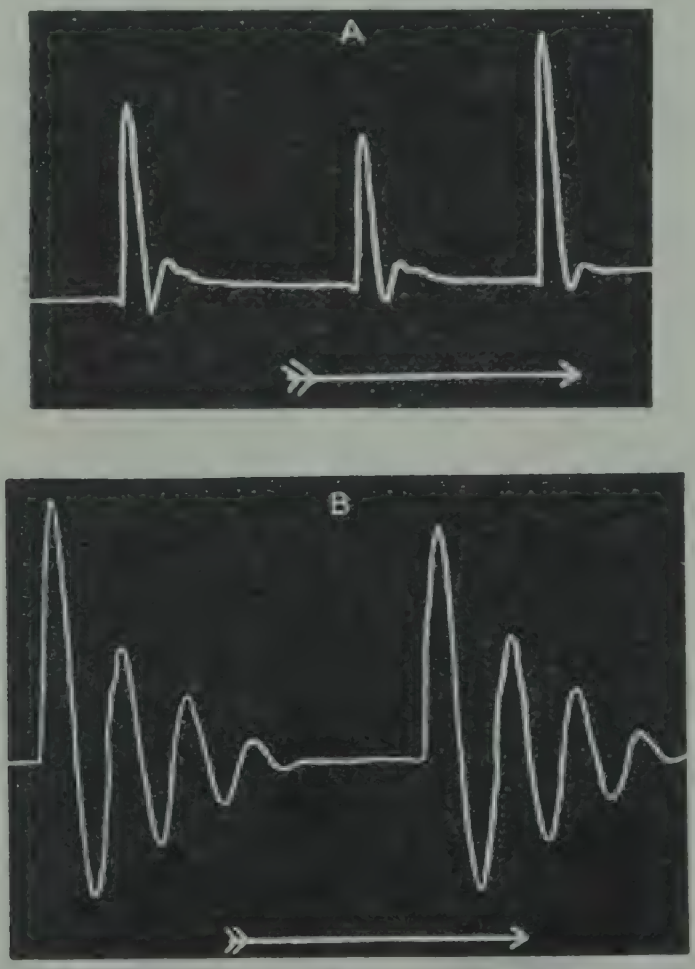


Fig. 227. Records of Knee-jerks. Normal jerk above. Jerk affected by cerebellar injury below. (Gordon Holmes, *Brain*, 1918, **40,** 461.)

Neocerebellar Destruction. A definite syndrome results from loss of the lateral lobe or lobes of the cerebellum, many cases of which, resulting from bullet wounds, were studied by Gordon Holmes. It is characterised by asthenia, definite hypotonia, disorders of voluntary movement, abnormal postures, and tremor on the same side as the lesion. These cause characteristic clinical signs, to many of which particular names have been given.

---

* Denny Brown *et al.* *Proc. R. S.*, 1929, **104B**, 518.

† Miller and Laughton. *Arch. Neurol. Psych.*, 1928, **19**, 47; *Proc. R. S.*, 1928, **103B**, 575.

(1) *Hypotonia.* The muscles on the affected side are flabby, and although knee-jerks are present, owing to depression of the normal plastic tone the jerk is of a lax and swinging character (Fig. 227). Hypotonia is probably the fundamental defect in neocerebellar lesions, and after removal of the whole cerebellum is also the dominant feature in some species of animals.

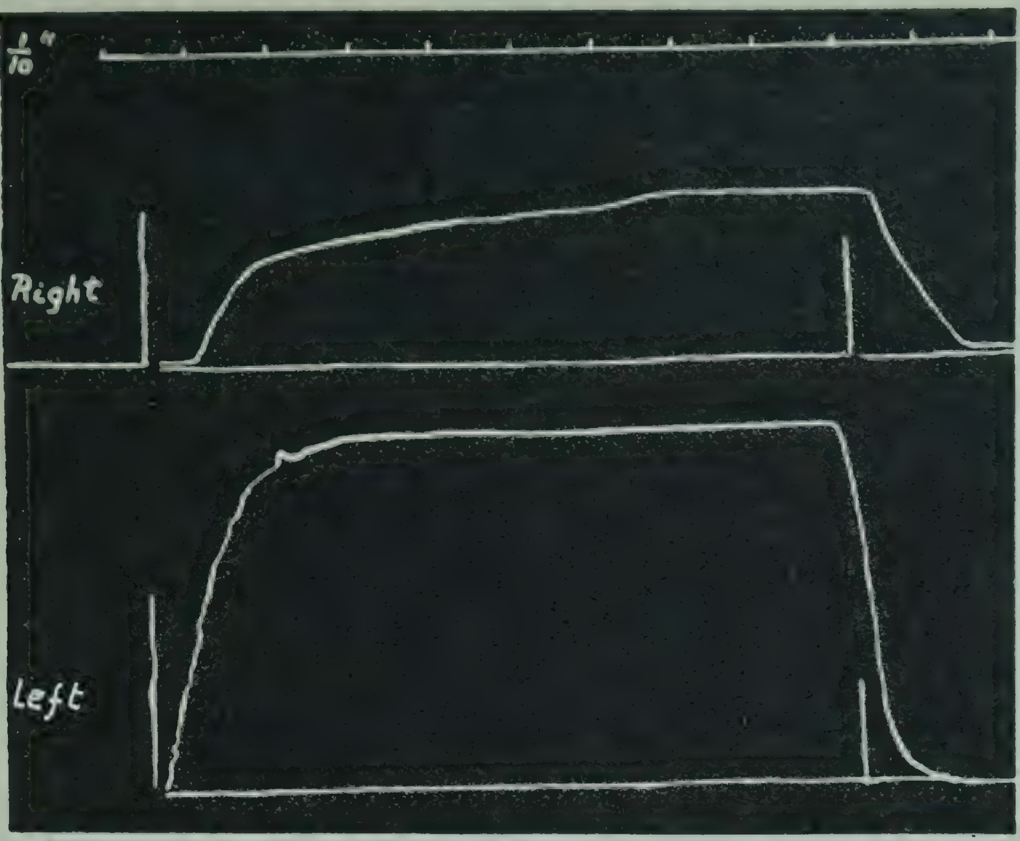


FIG. 228. Illustrating the Characters of Voluntary Contractions in Cerebellar Disease, which affected the Right Side of the Cerebellum. Slower and weaker contractions on the same side. (GORDON HOLMES. *Brain*, 1918, **40**, 461.)

(2) *Disorders of Voluntary Movement.* Voluntary contractions are weak (*asthenia*), slow in onset and ending, and easily show fatigue (Fig. 228). They also show errors in range and force (*dysmetria*). Hence rapidly alternating movements, such as pronation and supination, are difficult to perform (*adiadochokinesis*) (Fig. 229), and co-ordination of movements is poor (*ataxia*), because of improper association of the agonist and antagonist muscles concerned (*asynergia*). In order to compensate for these disabilities the patient often performs a movement in stages, as though " by numbers " (*decomposition of movement*). Speech disorders are sometimes present and are probably due to loss of delicacy and precision in muscular movements.

(3) *Abnormal Postures.* Because of interference with proprioceptive impulses, abnormal attitudes of the arms or legs may result : thus the arm held forwards deviates outwards on the affected side if the eyes are closed (*Barany's pointing test*). There is also a tendency to fall or deviate towards the injured side ; hence the gait zigzags. Nystagmus is common, (slow component towards side of lesion), but vertigo rare.

(4) *Tremor,* of a coarse character, appears as soon as a muscle or limb begins to suffer fatigue, but is only conspicuous if the deep nuclei are damaged.

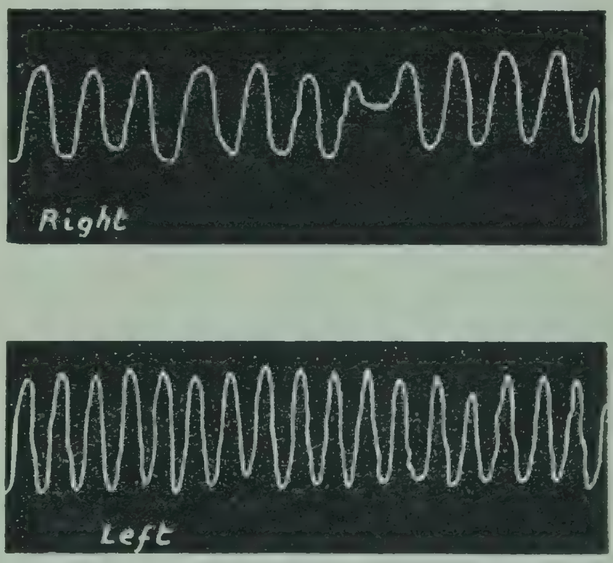


FIG. 229. Adiadochokinesis in Right-sided Cerebellar Lesion. Record of rapid supination and pronation movements of arm on injured side and on normal side (lower). (GORDON HOLMES. *Brain*, 1918, **40**, 461.)

There is rapid compensation for the effects of the lesion, and it is on the whole remarkable how small the effect of a large cerebellar loss may finally be.

CONCLUSIONS REGARDING THE FUNCTIONS OF THE CEREBELLUM

Three points seem clear about the functions of the cerebellum. First, that it is in no way connected with sensation ; second, that it has important relations to the maintenance of postural tonus ; and third, that it is concerned with the attainment of precision as regards force, timing and direction in *voluntary* muscular movements.

The phylogeny of the organ, its connections, and the effects of partial ablations all point to a difference of function as between the palæocerebellum and the neocerebellum. The former appears most clearly to be concerned in the fine adjustment and integration of the unconscious afferent impulses which play a part in the maintenance of the proper distribution of postural tonus. It is essentially proprioceptive, and involved in the co-ordination of the reflexes of muscular tonus in its relation to equilibrium maintenance. It probably, among other actions, serves to moderate the proprioceptive reflexes concerned in the maintenance of equilibrium. In man this function is a subsidiary one. The flocculo-nodular lobe is probably the control region involved in motion sickness.*

It is perhaps not possible to draw a perfectly definite line separating off the functions of the neocerebellum from those of the palæocerebellum, because normal voluntary movements are played off on a background of postural tonus, but it is evident that the neocerebellum, the development of which is phylogenetically parallel to that of the cortex, has important functions with respect to the execution of voluntary movements. We may suppose from the evidence that has been reviewed that it serves in some way to organise and, where necessary, to fortify the impulses from the cerebral cortex, and so to fit them into the existing background of postural tonus as to give precise, smooth and sufficiently powerful contractions of the various muscle groups in their correct relations to each other. Hence, in some way at present unknown, it is superadded to, and seems to exercise a kind of final editorial supervision over, the motor impulses emerging from the cerebral cortex. " By virtue of the proprioceptive impulses which it receives from all peripheral structures concerned in posture and movement, the cerebellum is in a position to send to the forebrain an integrated influx which may serve in the control and distribution of muscular tone " (Gordon Holmes, 1939). Local representation of muscular areas on its cortex is at most only vaguely indicated and indeed a study of its histological structure leads to the conclusion that its functions are generalised over its surface.

For Reference

ASS. RES. NERV. MENT. DIS. (1926). The Cerebellum. (Baltimore : Williams and Wilkins Co.)

BREMER (1935). Le cervelet, in Roger and Binet's *Traite de Physiologie*, **10**, 39–134.

HOLMES, GORDON (1918). The Symptoms of Acute Cerebellar Injuries due to Gunshot Injuries. *Brain*, **40**, 461.

HOLMES, GORDON (1922). Croonian Lectures, R.C.P. *Lancet*, **100** (1) 1177, 1231 ; (2) 59, 111.

HOLMES, GORDON (1939). The Cerebellum of Man. *Brain*, **62**, 1.

LUCIANI (1915). Human Physiology, III, **8**, 419. (London : Macmillan.)

MILLER (1926). The Physiology of the Cerebellum. *Physiol. Rev.*, **6**, 124.

---

* BARD, *et al.*, *Fed. Proc.* 1947, **6**, 72.

CHAPTER XXIII

THE AUTONOMIC SYSTEM: THE HYPOTHALAMUS: HUMORAL TRANSMISSION

I. THE AUTONOMIC SYSTEM

The autonomic nervous system includes those arrangements of efferent and afferent nerve fibres which are primarily concerned with the regulation of the internal processes of the body. Its efferent fibres are distinguished from the somatic efferents supplying skeletal muscle by having in the course of each ultimate fibre one synaptic connection outside the central nervous system, viz. in one of the various ganglia of the sympathetic and parasympathetic systems. Its afferent fibres are widespread and subserve many important reflex functions as well as being concerned in the appreciation of visceral pain.

The autonomic system has connections with, and tracts and centres in, the central nervous system proper, and according to the manner of the connections with the central nervous system, the efferent parts of the autonomic system are subdivided into two sub-systems of distinct, or even contrasting, function, viz. the *sympathetic* and *parasympathetic* systems.*

We may therefore consider the autonomic system under these divisions :

(1) Sympathetic—or thoracico-lumbar } Efferent
(2) Parasympathetic—or cranio-sacral } Efferent
(3) Afferent autonomic fibres.

(1) **THE SYMPATHETIC SYSTEM** (Fig. 230) consists of the *lateral chain* of sympathetic ganglia, various *collateral ganglia* and plexuses, and many scattered *terminal ganglia*. The whole system originates from neurones limited to the intermedio-lateral cell columns of the thoracic and upper lumbar levels of the spinal cord, though the sympathetic chain extends from the upper cervical level to the coccyx. The cells of the sympathetic ganglia are all multipolar.

The sympathetic chain consists of bundles of nerve fibres, medullated and non-medullated, which connect together longitudinally the "*lateral*" *ganglia* lying on each side of the vertebral column, there being as a rule one ganglion to each spinal nerve root. But in the cervical region these ganglia are condensed into two or three, the *superior, middle and inferior cervical ganglia*, united by the cervical sympathetic trunk ; and the upper three or four thoracic ganglia on each side are condensed, sometimes, as in cat and dog, together with the inferior cervical ganglia, to form the "*stellate*" *ganglion*.

Details differ from one species, or even individual, to another. In man the inferior cervical ganglion is often fused with the first thoracic ganglion, and then corresponds to the stellate ganglion in cat or dog. The lowest ganglion of the cervical vagosympathetic trunk, and which lies on the oral side of the annulus of Vieussen's, is most commonly now regarded as the middle cervical ganglion.†

At the bottom of the chain there is only one coccygeal ganglion for the coccygeal vertebræ.

The (lateral) ganglia of the sympathetic chain are connected with all the spinal nerves, just after they have given off their posterior divisions, by means of the *rami communicantes*. These rami communicantes are of two kinds :

---

* Langley. *J. Physiol.*, 1905, **33,** 403 ; 1911, **43,** 173.

† Van den Broek. *Morphol. Jahrb.*, 1908, **37,** 202 ; Nonidez, *Amer. J. Anat.*, 1939, **65,** 361.

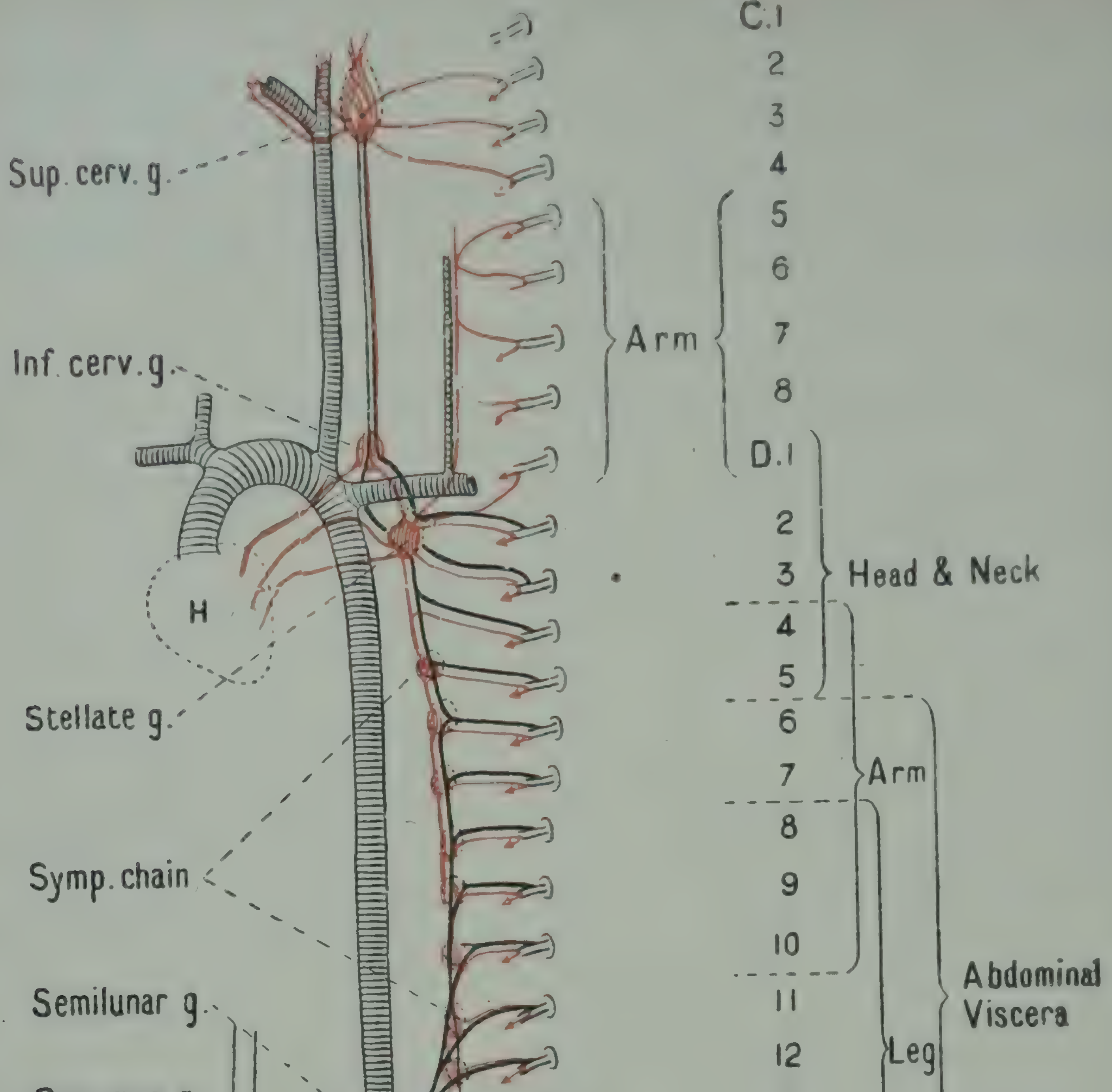

C.1
2
3
4
5
6
7
8
D.1
2
3
4
5
6
7
8
9
10
11
12
13
L.1
2
3
4
5
S.1
2
3
Arm
Head & Neck
Arm
Abdominal Viscera
Leg
Leg
Sup. cerv. g.
Inf. cerv. g.
H
Stellate g.
Symp. chain
Semilunar g.
Sup. mes. g.
S
I
Inf. mes. g.
C
Pelvic g.
B
Hypogastric n.
Pelv. visc. n.

FIG. 230. Diagrammatic Representation of the Distribution of the Sympathetic and Sacral Autonomic Systems.

The black lines represent the medullated pre-ganglionic fibres, such as those making up the white rami communicantes, while the post-ganglionic fibres are printed in red. On the extreme right of the figure is indicated the general distribution of the white rami arising from the several nerve roots, while the double brackets point to the nerve roots making up the limb plexuses. (The sacral autonomic fibres arise from S2 and 3, and are parasympathetic. The ganglion marked "inf. cerv. g." is, in most animals, now called the middle cervical ganglion, the inf. cervical g. being considered as fused into the stellate g.)

H. Heart.
S. Stomach.
I. Small intestine.
C. Colon.
B. Bladder.

white rami consisting of small medullated fibres, and *grey rami* composed almost exclusively of non-medullated nerves. It was shown by Gaskell that the white rami are formed by fibres which have their origin in the spinal cord and perhaps also in the posterior root ganglia ; whereas the grey rami represent fibres which, arising in the sympathetic ganglia, run back to join the spinal nerves (Fig. 231). The outflow represented by the white rami is limited to a distinct region of the cord, viz. from the first thoracic to the third or fourth lumbar nerve roots ; whereas the grey rami pass back from the sympathetic chain to all the spinal nerve roots. Stimulation of a limited number of white rami produces all the effects that can be evoked by stimulation of the grey rami, showing that the impulses leaving the cord pass upwards and downwards in the sympathetic system and are broken somewhere in their course, being transferred to a fresh relay which, by means of non-medullated nerves, carries them on to their destination.

Every fibre of the sympathetic system at some point of its course forms one synapse with a nerve cell, and Langley showed that this is the only cell-break in the fibre. This applies not only to the sympathetic fibres, but also to the fibres of the parasympathetic. Each fibre path therefore can be regarded as made up of two sections—a *pre-ganglionic* fibre arising in the spinal cord and passing out to a ganglion as a fine medullated nerve fibre, and a *post-ganglionic* fibre arising in this ganglion and continued as a non-medullated fibre, to its peripheral distribution.

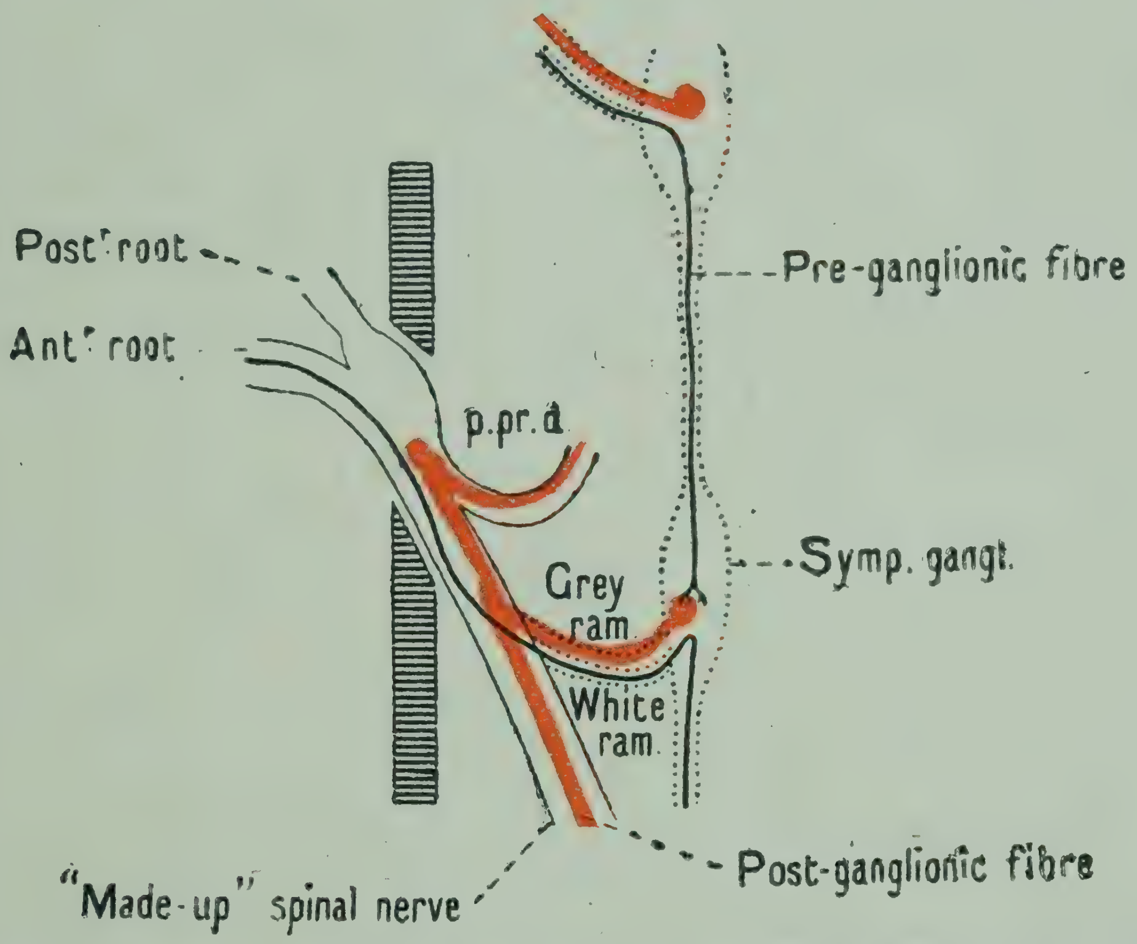


Fig. 231. Diagram to show the manner in which a Spinal Nerve is completed by the entry of a Grey Ramus, containing Fibres derived from Cells in the Sympathetic Chain. (After Langley.)

p.pr.d. **Posterior primary division. (The post-ganglionic fibres are represented in red.)**

This transference from one system to another involves the passage across a synapse and a nerve cell. The situation of this synapse may be easily revealed by utilising the action of nicotine, first studied by Langley.* If nicotine be applied to a sympathetic ganglion, it first stimulates and then paralyses any synapses which may lie in the ganglion. Intravenous injection of nicotine therefore causes a primary general excitation of all visceral ganglion cells. There is an enormous rise of blood pressure, which may be accompanied by other sympathetic effects, such as dilatation of the pupil, erection of the hairs, and so on. This rise rapidly passes off, and it is then found impossible to evoke any reflex visceral effects or any contraction, *e.g.* of the blood vessels by stimulation of the spinal cord; the passage of the impulses is blocked in every one of the visceral ganglia. By observing the effects of stimulation of a nerve before it enters a ganglion and then painting the ganglion with nicotine, and again trying the effects of excitation, it is easy to determine whether the nerve fibres which were excited in the first case form any synapses with the nerve cells of that ganglion.

---

* Langley. *J. Physiol.*, 1890, **11**, 123.

The relationships of the white and grey rami are strikingly illustrated in the case of the pilomotor fibres. These, in the cat, arise from the cord by the anterior roots from the fourth thoracic to the third lumbar inclusive. Passing by the white rami to the sympathetic system, they travel upwards and downwards and end by arborisations in the various ganglia of the main chain. From the cells of each ganglion a fresh relay of fibres starts, which runs as a bundle of non-medullated nerves (the grey ramus) to the corresponding spinal nerve, with which it is distributed to its peripheral destination. Each grey ramus causes erection of the hairs of one spinal segment, whereas stimulation of one white ramus causes erection in three or four segments, showing a distribution of the fibres of the white ramus to the cells in several successive ganglia. We thus see that, in speaking of the sympathetic functions of a spinal nerve root, we must clearly distinguish whether we mean the root as it arises from the spinal cord, in which case its visceral functions will include those of its white ramus, or whether we mean the made-up or complete spinal nerve after it has received its grey ramus (Fig. 231). In the latter case the visceral functions of the root will be more restricted than in the former case, and will have a different distribution.

The Collateral Ganglia. In the abdomen is a second system of ganglia, in special connection with the abdominal viscera, lying in front of the aorta and surrounding the origins of the large arteries to the alimentary canal. These are the semi-lunar, or cœliac, ganglia and the superior mesenteric and inferior mesenteric ganglia.

These ganglia are entangled in a fine feltwork of nerve fibres, which covers all the large arteries to form the sympathetic plexuses of the abdomen, *e.g.* the cœliac plexus, the mesenteric plexuses, hypogastric plexuses, etc. The outflow to these ganglia consists mainly of pre-ganglionic fibres, which have passed through the sympathetic chain without forming any synaptic connections there, and emerge from that chain as *the splanchnic nerves.* With them are also a few post-ganglionic fibres (Fig. 232).

The three splanchnic nerves originate from the fourth to the eleventh thoracic spinal roots, and pass to the cœliac and superior mesenteric ganglia. Small branches from the lumbar roots pass in a similar way to both the mesenteric ganglia. The majority of the splanchnic fibres form their synaptic connections in the collateral ganglia (Fig. 232). From the abdominal plexuses, which are largely non-medullated, fibres pass to all the abdominal viscera. There are also some fibres of pre-ganglionic type which pass directly to the viscera, where they form synapses with ganglion cells in their walls.

Distribution of the Sympathetic Fibres. The post-ganglionic fibres of the sympathetic system are distributed to their destinations in one of two main ways, viz. : with the spinal nerves, or by travelling along with the small arteries. The fibres of the grey rami are of the first type, the remainder are of the second.

There are grey rami to all the spinal nerve roots; this means that the grey rami may contain fibres derived from white rami at other levels, since the white rami are confined to the thoracic and upper lumbar roots. The pre-ganglionic fibre of the white ramus passes to the sympathetic chain and may travel up or down in it for one or more segments, giving off branches on the way, but ultimately it ends by synapse with many cells in one or more of the lateral ganglia. An impulse emerging from the cord by a pre-ganglionic fibre has therefore ample opportunity to affect many neurones of the chain, and to be dispersed along many differently distributed post-ganglionic fibres. Reflex responses of the sympathetic system therefore involve a wide and diffuse field of effectors.

(1) The "post-ganglionic" axons of the cells of the lateral ganglia enter the grey rami (often after some further travel up or down in the chain), and are then distributed to the region supplied by that particular spinal nerve root; these areas are the skin of the limbs and trunk, the structures supplied being the arterioles and capillaries (vaso-constrictor), the sweat glands (secretory), the hair muscles or *arrectores pilorum* (motor, causing erection of hairs): fibres also pass to the blood vessels of muscle, to which they are mainly vaso-dilator. The fibres for the *upper limb* leave the cord by the white rami of the upper thoracic roots, and have cell stations in the stellate and middle cervical ganglia. Those for the *lower limb* are derived from the lower thoracic and upper lumbar roots. Those for the trunk areas come from the middle regions of the outflow from the cord.

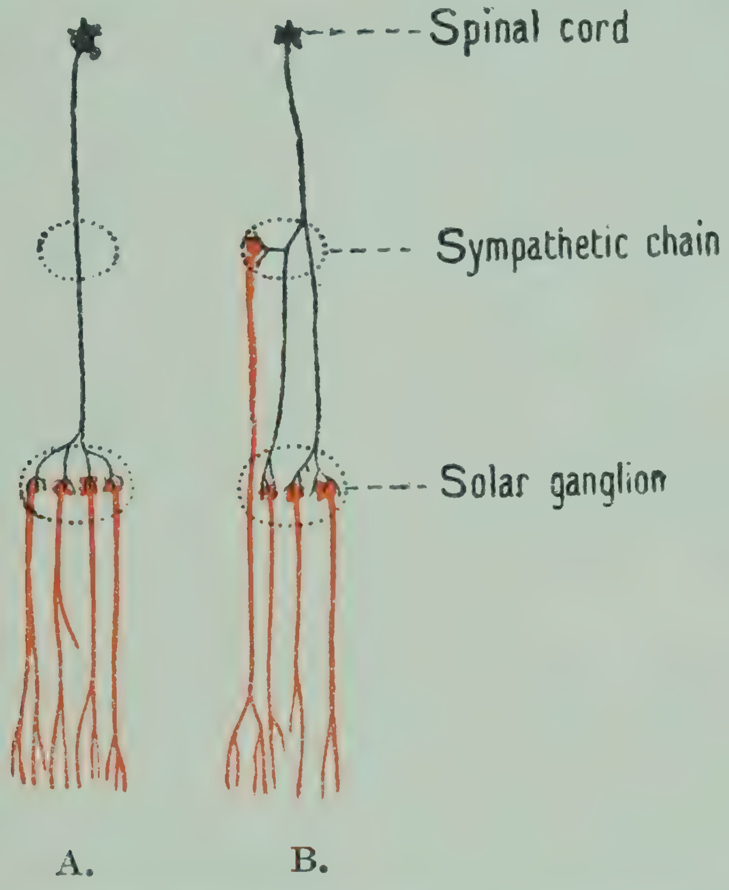


Fig. 232. Figure to show the probable Mode of Connection of the Fibres of the Splanchnic Nerve with Nerve Cells. (After Langley.)

A. Usual type, all the fibres passing through the lateral chain to end in the collateral ganglia of the solar plexus.

B. Alternative condition, in which a small minority of the fibres have their cell stations in the sympathetic chain.

The pre-ganglionic fibres are black, the post-ganglionic red.

(2) The remaining fibres of the sympathetic system are distributed by passing along the walls of the arteries, and most of the fibres to the skin of the head and neck are also of this type.

(i.) *The Head and Neck.* These fibres arise from the first and second thoracic nerves, ascend in the cervical sympathetic and have cell stations in the superior cervical ganglion, from which post-ganglionic fibres travel with the blood vessels, to supply all the structures of the head and neck. (*a*) The fibres to the *skin* supply sweat glands, arrectores pilorum muscles, and small blood vessels. (*b*) The *salivary*, lachrymal, and buccal glands receive fibres that are secretory and vaso-constrictor. (*c*) The *brain* receives vaso-constrictor fibres. (*d*) The *eye* receives motor fibres to the *dilator pupillæ* muscle, to the *nictitating membrane* (when present), to the plain muscle of the orbit (*Müller's muscle*), and to some plain muscle in the upper lid. Stimulation of the upper end of the cut cervical sympathetic therefore causes vaso-constriction in head and neck vessels, dilatation of the pupil, retraction of the nictitating membrane and upper lid, and some protrusion of the eyeball, while section, or excision of the superior cervical ganglion, gives the opposite effects. (*e*) The *thyroid* also receives fibres, the function of which is uncertain.

(ii.) *The thoracic organs* are supplied from the second to fourth thoracic nerves, with their cell stations situated in the stellate ganglion, and, in man, also in the inferior, and middle, and some even in the superior cervical ganglia. Post-ganglionic fibres are distributed with the cardiac, pulmonary and other plexuses. Those to the *heart* are accelerator and also dilate the coronary vessels ; those to the *lungs* are broncho-dilator and vaso-constrictor.

(iii.) *The abdominal organs* derive their pre-ganglionic fibres from the thoracic roots below the fourth or fifth, the fibres passing through the sympathetic chain to form the *splanchnic nerves.* These pass to the various collateral ganglia, the cœliac, and superior and inferior mesenteric ganglia and the pelvic or hypogastric ganglion, where the synapses are situated. Post-ganglionic fibres from these ganglia then travel along the blood vessels. These fibres, together with some pre-ganglionic fibres and numerous scattered small ganglia, form elaborate plexuses along all the main arteries of the abdomen, and are named after these arteries, *e.g.* renal plexus, etc. The *hypogastric nerves*, derived from the lower levels of splanchnic outflow, mainly relay in the hypogastric ganglion, the arrangement in man being different in detail from that in lower mammals.

All the abdominal and pelvic viscera are supplied with sympathetic fibres. (*a*) To the stomach and intestines, with the exception of their sphincters, the fibres are vaso-constrictor and inhibitory of movement. The sphincters are supplied with motor fibres, so that a result of sympathetic stimulation is a retention of contents in the alimentary canal. (*b*) The bladder is also supplied by inhibitory fibres, but its sphincter with motor fibres. (*c*) The liver is supplied with fibres that constrict its blood vessels and also cause a discharge of glucose from the liver into the blood. (*d*) Weak contraction of the gall-bladder may also be caused by the sympathetic.

(*e*) To the spleen the sympathetic sends motor fibres to its plain muscle, and the organ diminishes considerably in size owing to expulsion of contained blood, when the sympathetic fibres of the splenic nerve are stimulated.

(*f*) Lastly, the splanchnic also sends an important branch to the suprarenal body, which discharges adrenaline into the blood stream when the splanchnics are stimulated. These *fibres that go to the suprarenal are pre-ganglionic fibres*, and instead of forming synapses with ganglion cells, they end *directly in relation to the cells of the suprarenal medulla.* It is interesting to note, however, that these cells are actually developed from the same embryological structures as form the sympathetic ganglion cells—a fact of some significance, as we shall see.

FUNCTIONS OF THE AUTONOMIC GANGLIA

These ganglia consist of a mass of nerve cells embedded in connective tissue, each cell being surrounded by a special capsule. The nerve cells are multipolar, each cell possessing one axon and several dendrites. The dendrites end in little arborisations round adjacent cells.

Since the main nervous system is characterised by the possession of nerve cells, it was formerly thought that any collections of nerve cells must partake of the co-ordinating and reflex functions of the central nervous system, *i.e.* must act as local nervous centres. Though theoretically conceivable, all efforts have failed to prove the existence of such a function, and we must conclude that the sole use of these ganglia is to serve as distributing centres. Each pre-ganglionic fibre divides, and the branches arborise round several cells (Fig. 232). In the superior cervical sympathetic the ratio of pre- to post-ganglionic fibres is 1 : 32.* A complex of fibres and cells is found in mammals in the plexuses of Auerbach and Meissner, which do in fact act as local nerve centres for the intestine. But all such mechanism is wanting in the sympathetic ganglia.

There is, however, a certain group of phenomena which can be elicited in sympathetic ganglia, and which were termed by Langley and Anderson † *axon reflexes.* If, for instance, in the cat we divide all the nerves going from the sympathetic chain to the inferior mesenteric ganglion, leaving the bladder connected with this by the hypogastric nerves, and then after dividing the left nerve stimulate its central end, we obtain a contraction of the right half of the bladder. This effect is abolished by painting the inferior mesenteric ganglion with nicotine, showing that the activity of the cells of this ganglion is involved in the process. It has been shown, however, by Langley and Anderson that this is not a true reflex. A pre-ganglionic fibre arriving at the inferior mesenteric ganglion branches, one branch ending round the cells of the ganglion, while the other branch passes down in one hypogastric nerve to a cell situated near the base of the bladder (Fig. 233). When, therefore, we stimulate the central end of the divided nerve, we are stimulating some pre-ganglionic fibres, and the excitation spreads up to the point of junction of the two branches and then down the other branch to excite the cell in the inferior mesenteric ganglion. We thus obtain an apparent motor reflex by stimulation of a nerve which is itself motor. Similar pseudo-reflexes can be obtained for other viscera, *e.g.* the colon, *viâ* the hypogastric, as for the bladder, and also along the abdominal chain on the pilomotor nerves,‡ but furnish no grounds for ascribing to peripheral ganglia the property of reflex centres.

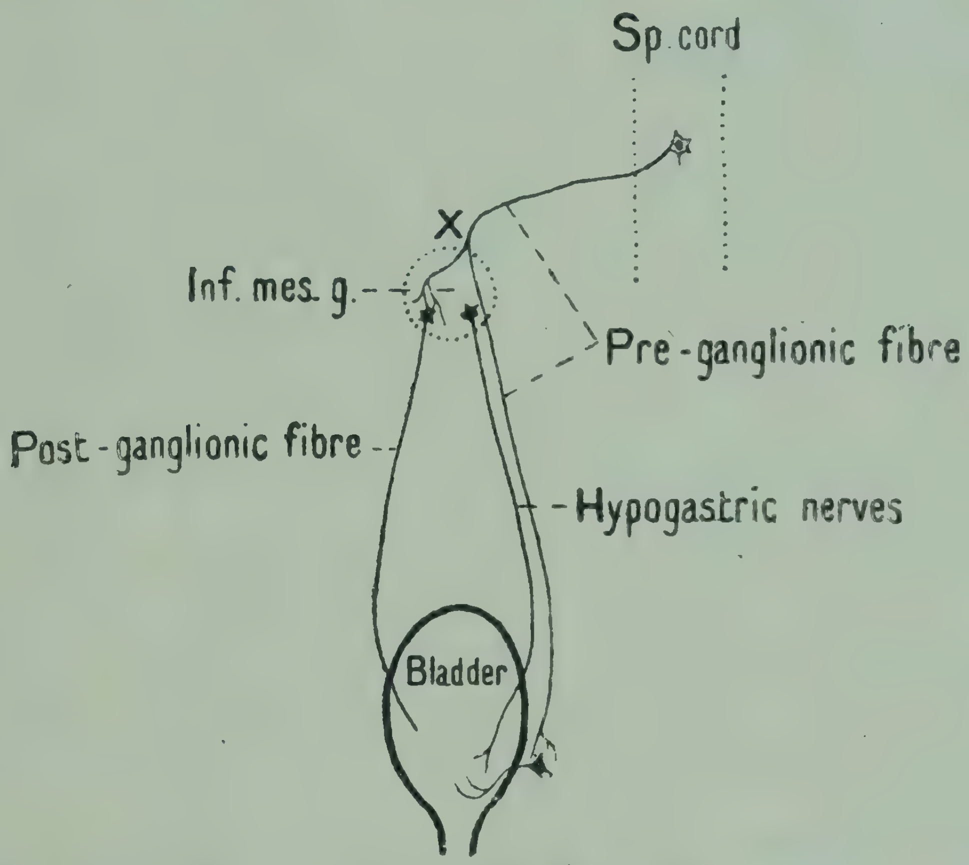


FIG. 233. Diagram to illustrate Langley and Anderson's explanation of the Hypogastric Reflex as an Axon Reflex.

The branching of the axon where the propagation or "reflexion" takes place is at X.

Another feature of sympathetic ganglia is that the effects of pre-ganglionic stimulation may be a prolonged state of excitation. This is possibly due to the presence of re-exciting circuits.§

DISPENSABILITY OF THE SYMPATHETIC SYSTEM

Complete removal or inactivation of the sympathetic system was performed on a number of animals by Cannon,‖ by removal of the thoracic and abdominal chains. Animals thus operated have lived for years in good health and of normal appearance, under the sheltered conditions of laboratory life. There is, however, a greatly reduced capacity to do work. This is because the body no longer has at its disposal the numerous reserves, such as heart acceleration, liberation of adrenaline, redistribution of blood to the working muscles, dilatation of bronchioles, etc., which are made available during exercise by the intermediation of the sympathetic nervous system. During emotional excitation or exposure to cold, also, there is none of the response associated with sympathetic activity, such as erection of hairs, or rise of blood sugar, or contraction of the spleen. It is therefore difficult to prevent a rapid fall of body temperature when the animals are exposed to moderate degrees of cold, in spite of the fact that the animals shiver much more readily than do normal animals under similar conditions. The same difficulty of regulating body temperature is seen when the surroundings are hot, so that dangerous rises of temperature are likely to occur. It is clear, therefore, that the sympathetic nervous system, though not essential to the maintenance of life, when conditions are favourable, has as an important, if not, as its chief, function, the automatic adjustment of the manifold responses of the individual to the emergencies of an adverse environment.

---

* RANSON and BILLINGSLEY. *J. Comp. Neurol.*, 1918, **29**, 359.

† LANGLEY and ANDERSON. *J. Physiol.*, 1893, **16**, 410.

‡ LANGLEY and ANDERSON. *J. Physiol.*, 1900, **25**, 364.

§ BRONK *et al.* *Amer. J. Physiol.*, 1938, **122**, 1.

‖ CANNON, NEWTON, BRIGHT, MEMKIN and MOORE. *Amer. J. Physiol.*, 1929, **89**, 84.
Technique, *v.* :—
BACQ *et al.* *Arch. int. Pharmacodyn.*, 1934, **48**, 429.
VERNEY and VOGT. *Q.J.E. Phys.*, 1938, **28**, 253.

(2) THE PARASYMPATHETIC SYSTEM consists of the cranio-sacral outflow of the autonomic system, and is divisible into the mid-brain, bulbar and sacral portions. Its pre-ganglionic fibres pass near to their final destinations before terminating in synapses, which, as shown by the nicotine method, are situated in peripherally placed ganglia.

The mid-brain outflow is represented by fibres in the IIIrd cranial nerve, which arises from the group of oculo-motor nuclei lying ventral to the aqueduct of Sylvius. The pre-ganglionic fibres pass to the ciliary ganglion in the orbit, from which they are relayed by post-ganglionic fibres to the *constrictor pupillæ muscle* and the *ciliary muscle.* Stimulation of the IIIrd nerve leads to constriction of the pupil, and contraction of the ciliary muscle. (Other fibres of the IIIrd nerve are not autonomic, viz. the nerves, both afferent and efferent, to the extrinsic eye muscles.)

The bulbar outflow comprises some of the fibres in the VIIth, IXth and Xth cranial nerves. In the *VIIth nerve* the fibres pass to the spheno-palatine ganglion, where relays of fibres occur for supply of the *lachrymal gland* and structures in the pharynx, and other fibres pass *viâ* the *chorda tympani* to peripheral ganglia, for supply of the secretory and vasodilator fibres of the *submaxillary* and *sublingual* salivary glands. The *IXth nerve* similarly, *viâ* the otic ganglion, supplies the *parotid* gland and tongue with secretory and vasodilator fibres.

The *Xth cranial nerve* supplies most of the thoracic and abdominal viscera with parasympathetic fibres, which reach their destinations by way of various autonomic nerve plexuses, in which, as well as in the organs themselves, are placed the widely scattered ganglia on their course. The vagus fibres so distributed are inhibitory to the *heart*, vasoconstrictor to the heart vessels, constrictor to the *bronchi*, motor to the lower *œsophagus*, and augment movements of the *stomach* and *intestines*, with the exception of the sphincters, to which they are inhibitory. On the *stomach and pancreas* it has secretory functions, and also causes vasodilatation of the small blood vessels.

The sacral outflow leaves the spinal cord by the second and third sacral ventral roots, which form the *pelvic visceral nerves*, or *nervi erigentes.* These supply the pelvic viscera, the ganglia being mostly placed on their walls. The large intestine and rectum and bladder are supplied with motor fibres, and the external genitalia with vasodilator fibres. Stimulation of the pelvic visceral nerves causes erection of the penis as a result of the vasodilatation which follows.

COMPARISON BETWEEN SYMPATHETIC AND PARASYMPATHETIC

The various effects of the parasympathetic system are directed with greater precision of localisation to their effector organs than is the case with the sympathetic system. With the latter, a pre-ganglionic fibre makes connection with several ganglion cells which are all situated at a distance from the effector organs or tissues, so that one fibre may produce an effect over a wide area. The parasympathetic fibres, on the other hand, proceed as pre-ganglionic fibres until they are near to their destinations, so that there is much less diffusion of their effect.

Action of Drugs. The widespread effects of the sympathetic system are also ensured by the fact that in states of sympathetic excitation, *adrenaline* is liberated from the suprarenal body into the blood, branches from the splanchnic nerves being secretory to the medulla of the suprarenal. The injection of adrenaline into the circulation reproduces most of the effects of sympathetic stimulation all over the body, *e.g.* dilatation of pupil, widespread vasoconstriction, liberation of glucose from the liver, etc. By circulating all over the body, therefore, the adrenaline which is set free by splanchnic stimulation, causes a generalisation and prolongation of sympathetic effects in nearly all the organs that have a sympathetic supply. Thus, stimulation of the splanchnics not only causes relaxation of the tone of the stomach and intestines, but also acceleration of the heart and intense vasoconstriction of small vessels in the skin all over the body, even if the nerve supply to them has been cut.

Ergotoxine (ergotamine) paralyses the motor and secretory fibres of the sympathetic system, but has no effect on the inhibitory functions. If the drug is given, and the splanchnic then stimulated, or adrenaline injected, general vasodilatation occurs instead of constriction, so that the blood pressure falls. This is because there are sympathetic vasodilators, unaffected by the ergotoxine; normally their action is masked by that of the more powerful vasoconstrictors.

Dibenamine (n, n-dibenzyl-chloroethylamine)* has similar actions.

Cocaine greatly augments the effects of sympathetic stimulation or of adrenaline. *Ephedrine* has a similar effect, and it has been suggested by Gaddum that this is due to the fact that these substances protect the adrenaline from oxidation in the tissues.

Nicotine first excites and then paralyses all autonomic synapses, the sympathetic endings in the suprarenal medulla, and striated muscle motor end plates. *Curare* also paralyses synaptic transmission.

Muscarine (*and pilocarpine*) produce excitant effects on the structures supplied with post-ganglionic parasympathetic fibres, and also on sweat glands. All its actions are abolished by atropine.

Acetyl choline, when injected into the blood or applied to a tissue, reproduces all the effects of parasympathetic stimulation, and also causes secretion of sweat, although the sweat glands receive a sympathetic supply. It also has a nicotine-like action in stimulating certain skeletal muscles and autonomic ganglia. Its action is very transient, however, as it is rapidly destroyed in the blood and tissues by the action of an enzyme called choline esterase.

Eserine (di-isopropyl fluoro-phosphate (D.F.P.)), and certain alkyl polyphosphates† inhibit the action of this enzyme by becoming attached to its active group; hence, after eserine has been injected into an animal, acetyl choline then causes intense and prolonged effects, *e.g.* slowing or inhibition of the heart, secretion of saliva, etc. After eserine, the effects of parasympathetic stimulation are prolonged, and may also become generalised, *e.g.* stimulation of the vagus not only slows the heart, but also causes a secretion of saliva, etc.

---

* NICKERSON and GOODMAN, *J. Pharmacol.*, 1947, **89,** 167.

† BURGEN, *et al., Nature*, 1947, **160,** 760.

Atropine paralyses the actions of muscarine, and the muscarine-like effects, but not the nicotine-like actions of acetyl choline; it also renders parasympathetic stimulation ineffective.

For Reference

Darrow (1943). Physiological and Clinical Tests of Autonomic Function and Autonomic Balance. *Physiol. Rev.*, **23**, 1.
Gaskell (1916). The Involuntary Nervous System. (London: Longmans, Green.)
Henderson and Roepke (1937). Drugs affecting Parasympathetic Nerves. *Physiol. Rev.*, **17**, 373.
Kuntz (1946). The Autonomic Nervous System. (London: Baillière, Tindall and Cox.)
White (1935). The Autonomic Nervous System. (Macmillan.)

(3) AFFERENT FIBRES OF THE AUTONOMIC SYSTEM

There are abundant afferent fibres, many of them non-myelinated, in all the autonomic nerves. Thus, stimulation of the cut central end of the splanchnic nerve produces reflex effects such as a dilatation of the pupil, rise of blood pressure, etc. The autonomic afferents come from all parts of the body, and commence in receptors in the viscera and somatic structures.

Spinal autonomic afferents pass from the periphery in the splanchnic nerves, phrenics, etc., and some also with the various spinal nerves. They pass without interruption through the sympathetic ganglia, and white rami communicantes * to the dorsal roots, where their cell bodies are situated, as for

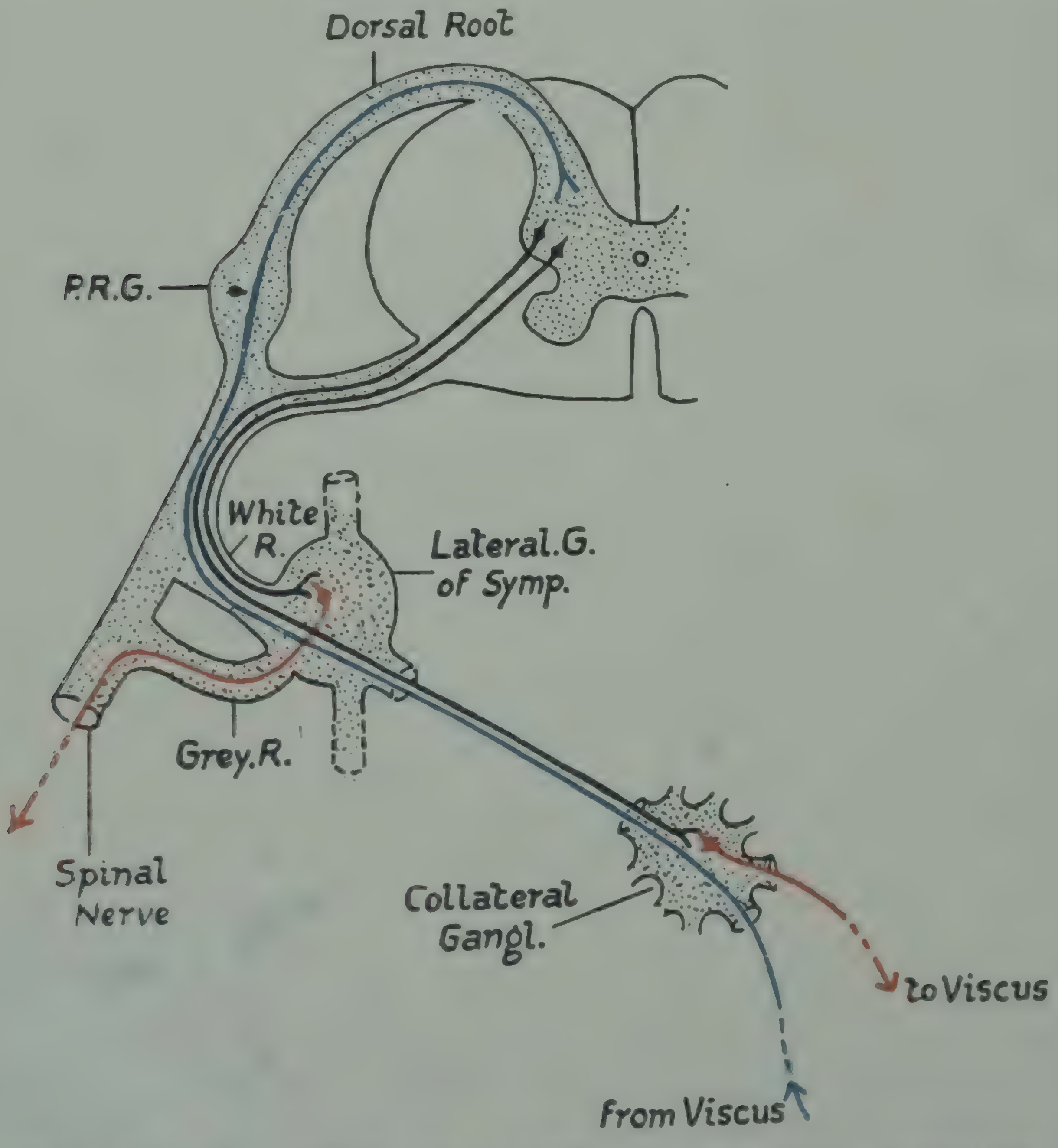


Fig. 234. Diagram showing the Afferent and Efferent Connections of the Sympathetic System. Afferent fibres blue, preganglionic efferents black, postganglionic efferents red.

---

* Ranson and Billingsley. *J. Comp. Neurol.*, 1918, **29**, 405.

the somatic afferents ; like these they enter the dorsal roots of the spinal cord (Fig. 234). The spinal roots concerned are all those that have white rami, viz. the thoracic and upper 3 lumbar and the 2nd to 4th sacral roots.

In the spinal cord the afferent autonomic fibres probably branch, much as the somatic afferent fibres do, and some terminate in relation to cells of the dorsal horn of grey matter. From here the further path may diffuse widely, by a series of relays, but sooner or later some of the collaterals form synapses with cells in the intermedio-lateral cell columns in the cord, from which the efferent sympathetic fibres emerge. How widely the efferent effects of afferent autonomic stimuli may spread is shown by the fact that on stretching of the intestine, dilatation of the pupil occurs.*

Visceral Pain. Afferent nerves from the viscera and from the parietal serous membranes are of importance in connection with visceral pain. To cutting, burning, or pinching, the viscera are quite insensitive, yet everyone knows that pain can originate in the viscera. Clinical experience and physiological experiments converge to the conclusion that pain originating in the hollow viscera is associated with, and probably due to, unduly powerful contractions of their muscular walls. Distension not accompanied by increased muscular contraction, as, *e.g.*, in paralytic ileus, gives rise to no pain.

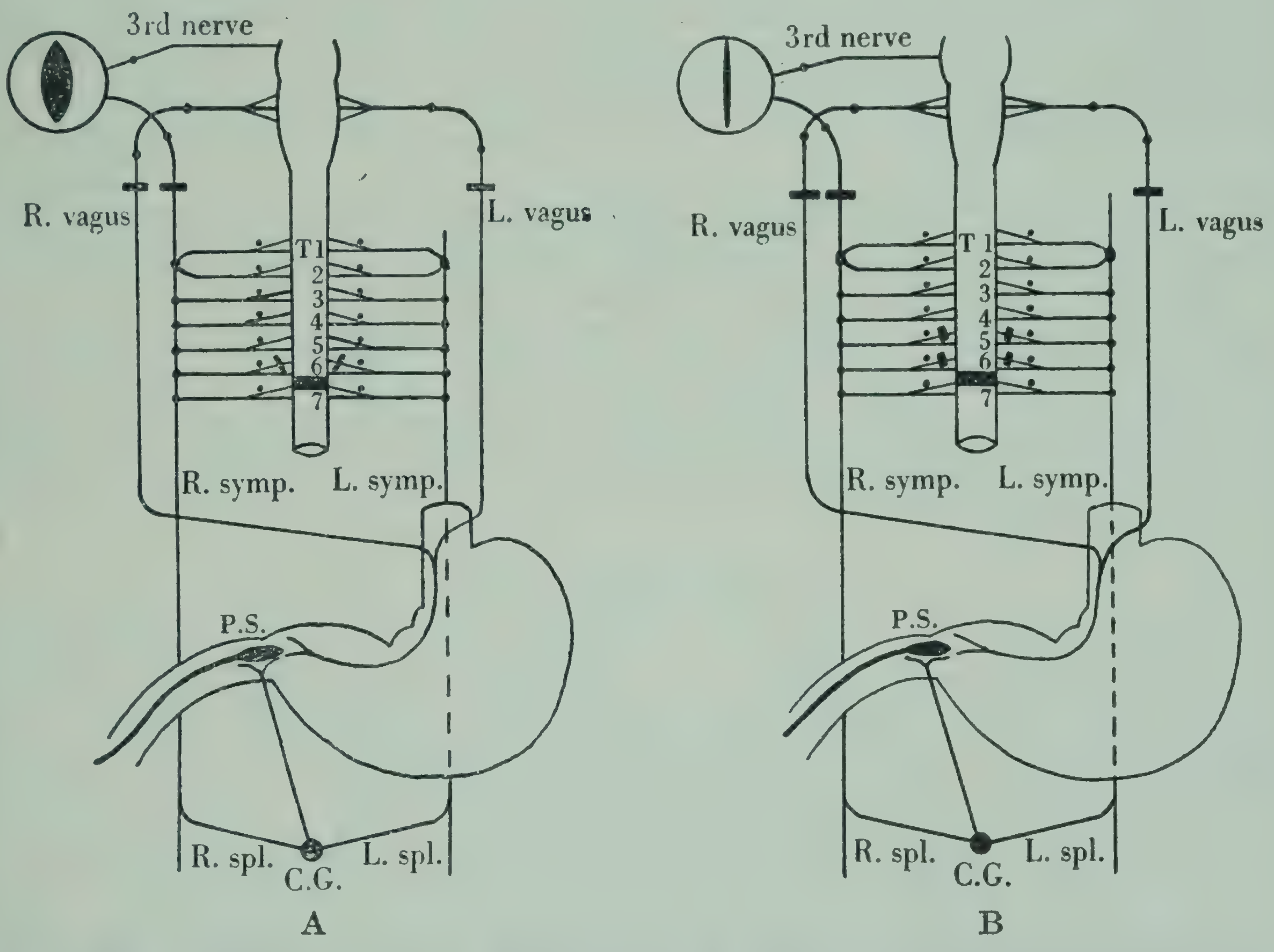


FIG. 235. Determination of Upper Limit of Entry of Afferents from Pyloric Sphincter. A. Cord cut between 6th and 7th Th. segments. R. and L. vagi and R. and L. cervical sympathetic nerves cut in the neck ; 6th Th. dorsal roots cut. Dilatation of pupil on distension of pyloric sphincter. B. On now cutting 5th Th. dorsal roots the response vanishes. C G = cœliac ganglion (McSWINEY and SUFFOLK. *J. Physiol.*, 1938, **93**, 104.)

Why the powerful contraction should cause pain is obscure : perhaps it may be related to the production of ischæmic pain (Chapter XXXVI) and so be referable to the local formation of a pain-stimulating factor.

Pain arising in the viscera is primarily felt in the region occupied by the viscus in early embryonic life, and in the case of the alimentary canal is around the mid-line.† Thus from stomach, duodenum or biliary passages it is felt in the epigastrium ; from small intestine or appendix in the umbilical region ;

---

* McDOWALL. *Q. J. Exp. Physiol.*, 1925, **16**, 177.

† BROWN. *B. M. J.*, 1942 (i), 543 ; *Lancet*, 1945, 10.

and from the colon, etc., in the hypogastrium. In addition to the above original sites of pain production, it may further arise from stimulation, especially from chemical excitation, of the somatic afferents of the parietal pleura or peritoneum, and is then often felt at a place with similar somatic innervation and accompanied by local pain, tenderness and muscular rigidity. For example, the later pain in appendicitis is felt in the right iliac fossa.

From the parietal pleura or peritoneum of the diaphragm, the pain is felt at the similarly-innervated shoulder tip (C. 3–5). The pain in angina pectoris, associated with over-distension of the heart and aorta radiates from the left chest, arm and shoulder (Th. 1–4).

When any part of the alimentary canal is distended by a balloon, a reflex rise of arterial pressure and dilatation of the pupil by inhibition of tone of the IIIrd nerve occurs.* By making sections of different parts of the central nervous system until this reaction disappears the path of entry and transit can be found. Fig. 235 shows that the afferents from the pyloric sphincter enter in and below the fifth thoracic root. The lower limit for pyloric sphincter afferents was the thirteenth thoracic root.

Cranial autonomic afferents are contained in the IIIrd, VIIth, IXth, Xth and XIth cranial nerves and perhaps in others. They have cell bodies in ganglia similar to the dorsal root ganglia, on the course of the nerves, and after entering the brain stem terminate in relation to cells of one of the afferent nuclei. Thus the afferents in the *vagus* are the axons of cells in the ganglion jugulare and ganglion nodosum. Their centrally running portions enter the medulla and branch, some of the branches ending in the dorsal vagal nucleus, in the floor of the IVth ventricle, in relation to cells the axons of which pass out in the vagus again as efferent autonomic fibres. This is a simple type of autonomic reflex arc, and many similar ones will be found in chapters dealing with the functions of the various viscera.

Another important example of an autonomic afferent is the nerve from the *carotid sinus*, a branch of the IXth cranial nerve. It conveys impulses resulting from stretching of the carotid artery and from changes in the composition of the blood therein, and important circulatory and respiratory reflexes are thereby brought about.

For Reference

Gask and Ross (1934). The Surgery of the Sympathetic Nervous System. (Baltimore: Wood.)

Ranson (1921). Afferent Paths for Visceral Reflexes. *Physiol. Rev.*, **1**, 477.

Hinsey (1934). The Anatomical Relations of the Sympathetic System to Visceral Sensation. *Proc. Assoc. Res. in Neur. and Mental Dis.*, **15**, 105.

THE INTEGRATION OF AUTONOMIC FUNCTIONS

It is evident that, in the main, the actions of the sympathetic and parasympathetic systems are opposed to one another, and many viscera have both supplies. Generally, the sympathetic system brings about changes that enable the organism to react to emergencies in such a way as to maintain a constant environment for its cells, or to help to protect itself by fight or flight, by bringing all its reserves into action, even at some cost. The parasympathetic system is similarly credited with more conservative functions, slowing the heart, reducing light to the eye, aiding digestion and generally serving a recuperative function.

It is perhaps more useful, however, to regard these two systems as both playing their parts in maintaining uniformity of conditions for the tissues of the body, by reason of their activities being balanced one against the other in such a way as to enable the viscera to work within a very wide range according to the varying requirements. The exact adjustment of body temperature or local blood supply, for example, requires a delicately balanced co-operation of both sympathetic and parasympathetic activities, and physiological enquiry is in fact very largely occupied in unravelling such co-ordinated activities, examples of which may be found in all branches of the subject.

---

* Irving and McSwiney. *J. Physiol.*, 1935, **85**, 4P.

It would appear that there are levels of integrative function in the autonomic system, as in the somatic system, and here also these become of increasing complexity as we go higher in the central nervous system.

There are autonomic vasomotor, pilomotor and sweat centres in the thoracic spinal cord, and centres controlling pelvic organs in the lumbar and sacral cord, through all of which simple reflexes may be obtained involving only the spinal cord, *e.g.* a rise of blood pressure on stimulating the central end of the splanchnic nerve in the spinal cat.

In the various afferent and efferent nuclei of the medulla and mid-brain we have not merely centres for the carrying out of special reflexes with which those efferent nuclei are concerned, *e.g.* the salivary nuclei in the lower pons and upper medulla, a vasomotor centre and a cardiac centre in the lower medulla, but also there are co-ordinating centres which exert an effect on the centres that lie lower down. Thus the medullary vasomotor centre exercises control on the spinal vasomotor centres, a centre in the mid-brain is concerned in co-ordinating the spinal reflexes of micturition, etc. More complex reflexes, *e.g.* swallowing, or salivation, as well as extensive circulatory and respiratory reflexes can be seen in the decerebrate cat, but this animal still lacks extensively co-ordinated reflexes, *e.g.* it cannot regulate its body temperature, or its blood sugar level.

When the hypothalamus is left behind, however, as in the thalamic animal, the body temperature and other autonomically regulated characters remain normal, and it is now generally agreed that it is in the hypothalamus that the highest refinements of autonomic adjustment are effected. There is in addition some degree of voluntary cortical control of autonomic function, *e.g.* in evacuation of the bladder, modification of breathing, etc., but this is superadded and does little to modify the integrated result of the hypothalamic functions.

II. THE HYPOTHALAMUS

The hypothalamus, which lies at the base of the diencephalon, below the thalamus, forms part of the wall and floor of the IIIrd ventricle. As seen from the base of the brain, it occupies the interpeduncular space and comprises the tuber cinereum and hypophysis anteriorly, and the corpora mammillaria posteriorly ; between the latter, and forming the floor of the IIIrd ventricle is the " central grey matter." Dorsally, it is separated from the thalamus by a fissure. The hypothalamus has an abundant blood supply.

The *hypophysis* is an endocrine organ, which will be further discussed in a later chapter ; it has important functional relations to the other hypothalamic structures.

Hypothalamic nuclei. Scattered throughout the hypothalamus are numerous nuclei and certain tracts. The nuclei and their connections are complex, and have not yet been fully worked out, but Le Gros Clark * has outlined a classification, as follows :—

(1) *Anterior Group*, above the optic chiasma.
    (*a*) The paraventricular nucleus, near the IIIrd ventricle.
    (*b*) The supra-optic nucleus, just above the optic chiasma.

(2) *Middle Group*, related to the infundibulum and tuber cinereum.
    (*a*) Ventro-medial nucleus.
    (*b*) Dorso-medial nucleus.
    (*c*) Lateral nuclei.
    (*d*) N. tuberalis.

---

* Le Gros Clark. *J. Anat.*, 1936, **70**, 203.

(3) *Posterior Group*, related to the mamillary bodies.
 (*a*) Posterior nucleus, above mammillary bodies.
 (*b*) Corpus mammillare nuclei.

Some of these nuclei are shown diagrammatically in Fig. 236.

Connections. There are many short connections, not yet fully mapped out, between the various hypothalamic nuclei, and with adjacent structures. There are also longer afferent and efferent connections.

Afferent fibres reach the hypothalamus from

(1) The olfactory tract. From the hippocampus *viâ* the column of the fornix to the mammillary bodies.

(2) Thalamus to posterior nucleus.

(3) Various parts of the body, *e.g.* from vagus, various ascending tracts in mid-brain, etc.

(4) Cortex, *viâ* median forebrain bundle, and unknown paths (not definitely proved).

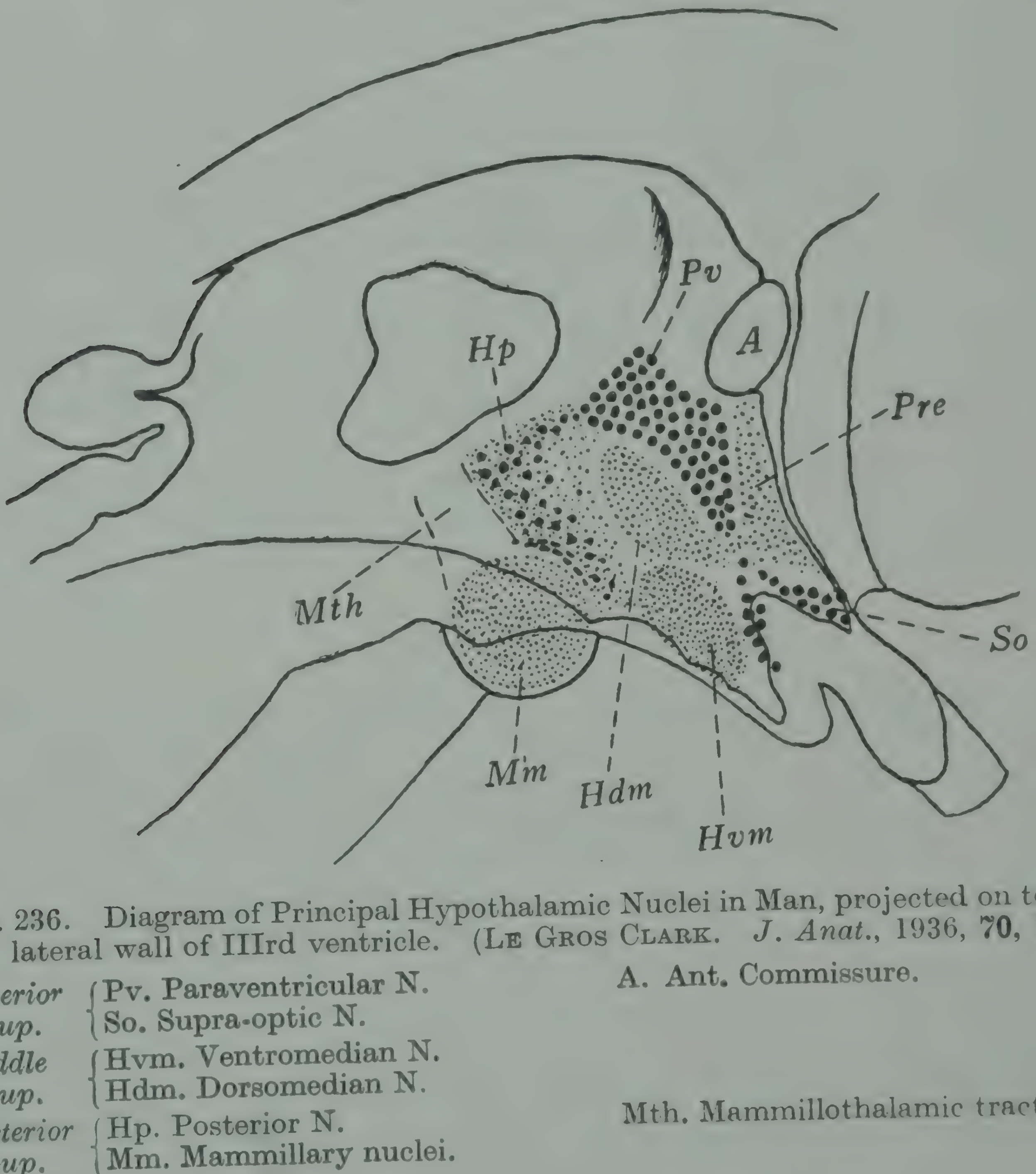


FIG. 236. Diagram of Principal Hypothalamic Nuclei in Man, projected on to the lateral wall of IIIrd ventricle. (LE GROS CLARK. *J. Anat.*, 1936, **70**, 204.)

Anterior Group. { Pv. Paraventricular N. / So. Supra-optic N.
Middle Group. { Hvm. Ventromedian N. / Hdm. Dorsomedian N.
Posterior Group. { Hp. Posterior N. / Mm. Mammillary nuclei.

A. Ant. Commissure.
Mth. Mammillothalamic tract.

Efferent connections from the hypothalamus pass to

(1) The anterior thalamus, from the mammillary bodies (the mammillothalamic tract is the most conspicuous tract in this region).

(2) The cortex, *viâ* the thalamus and by other routes.

(3) Autonomic nuclei of the brain stem and spinal cord (intermedio-lateral cell column). This is the most important of the descending efferent tracts : it probably starts in the lateral nuclei of the hypothalamus and leads to the reticular formation of the brain stem,* from which a series of relays passes t

---

* BEATTIE, BROW and LONG. *Proc. R. S.*, 1930, **106B**, 253.

the efferent neurones giving origin to the pre-ganglionic autonomic fibres. The connections are largely uncrossed, or if crossed, become so quite near their endings in the cord.*

(4) The hypophysis, by way of the supra-optico-hypophyseal and tubero-hypophyseal tracts which run in the stalk to reach the posterior lobe of the pituitary body (Fig. 237).

Functions of the Hypothalamus. *Stimulation.* It was found by Karplus and Kreidl in 1908 that on electrical stimulation of the hypothalamus there was widespread excitation of the sympathetic nervous system ; Beattie, Brow and Long † produced evidence that there were tracts connecting the hypothalamus with the spinal sympathetic centres ; and Beattie and Sheehan ‡ showed that both sympathetic and parasympathetic effects (*e.g.* reduction or increase of gastric motility) could be produced by suitable excitation, in the former case of the posterior nuclei, and in the latter of the tuber region. Detailed studies of the results of strictly localised excitation of known points in the hypothalamus have been carried out by Ranson and Magoun and their colleagues,§ and much precise information obtained.

In sum, the results have shown that general sympathetic responses, such as cardiac acceleration, rise of blood pressure, pupil dilatation, pilo-erection, hyperglycæmia, inhibition of bladder and intestinal tonus, resulted from stimulation of the posterior and lateral group of hypothalamic nuclei. Increase of respiration rate also occurred. Stimulation of the middle group of nuclei, in the tuber region, caused parasympathetic responses, *e.g.* bladder contraction, cardiac inhibition, and increased motility of the alimentary canal ; it also caused sweating. Anterior group stimuli also caused some parasympathetic effects of doubtful significance, and possibly caused changes in the water balance of the body. These last might be related to effects on the pituitary.

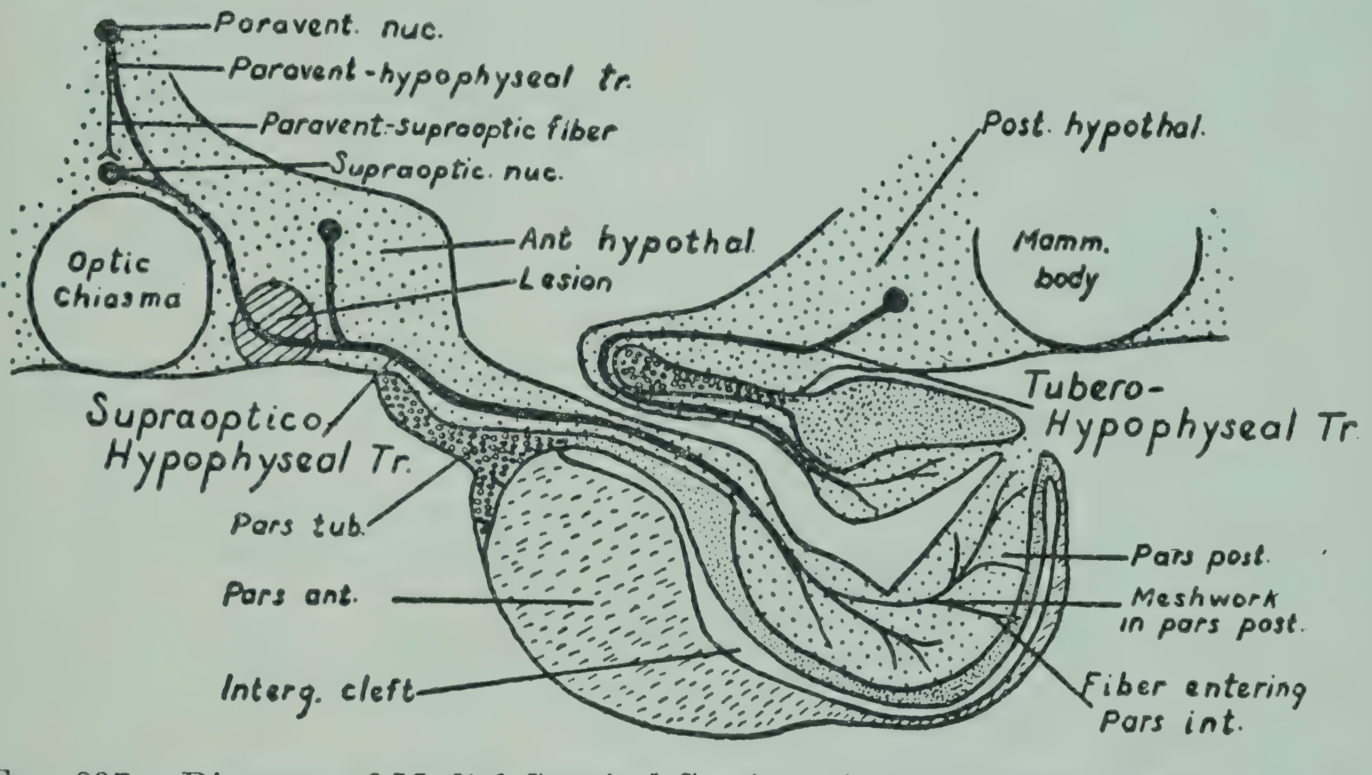

Paravent. nuc.
Paravent.-hypophyseal tr.
Paravent.-supraoptic fiber
Supraoptic. nuc.
Post. hypothal.
Optic Chiasma
Ant. hypothal.
Lesion
Mamm. body
Tubero-Hypophyseal Tr.
Supraoptico Hypophyseal Tr.
Pars tub.
Pars ant.
Pars post.
Meshwork in pars post.
Interg. cleft
Fiber entering Pars int.

FIG. 237. Diagram of Medial Saggital Section of Hypothalamic Region of Cat showing Nervous Connections of Posterior Lobe of Pituitary Body. (RANSON, Harvey Lecture, 1936–37.)

It has been shown (p. 312) that hypothalamic stimulation can cause facilitation of the lower motor neurones involved in spinal reflexes, and it has also been shown ‖ that facilitation and intensification of the effects of cortical motor stimulation can be brought about in a similar way. The path in this case, as traced by following strychnine action potentials seems to be from the hypothalamic nuclei to dorsomedial thalamic nuclei and thence to the cortex. Spinal afferent impressions can also enter this path by first reaching ventrolateral thalamic nuclei, and then → hypothalamus → dorsomedial thalamic nuclei → cortex.

---

* HARRISON, WANG and BERRY. *Amer. J. Physiol.*, 1939, **125**, 449.
† BEATTIE, BROW and LONG. *Proc. R. S.*, 1930, **106B**, 253.
‡ BEATTIE and SHEEHAN. *J. Physiol.*, 1934, **81**, 218.
§ RANSON. *Ergebnisse d. Physiol.*, 1939, **41**, 56.
‖ MURPHY and GELLHORN. *J. Neurophysiol.*, 1945, **8**, 339, 341, 431.

When the hypothalamus is stimulated in a conscious cat by electrodes fixed into the skull, all the symptoms of intense emotional excitement result, accompanied by powerful sympathetic effects, and it will be recalled that outbursts of "sham rage" with intense sympathetic excitation are also seen in decorticate cats, and are explained as release phenomena due to removal of cortical control.

Injection into the lateral ventricles of the brain of pituitary extracts, pilocarpine or acetyl choline, appears to stimulate certain hypothalamic nuclei, causing general sweating, vasodilatation, salivation, vomiting, etc.*

Lesions. Many remarkable effects have been found to be associated with lesions of the hypothalamus. They fall roughly into two groups, viz. those closely related to the functions of the pituitary body, and those with purely nervous effects. As regards the former group, it is still not clear where the functions of the pituitary body, which are mainly endocrine, cease and those of the hypothalamus begin, and many of the effects of lesions of the one are also found when damage is done to the other. They are closely related.

In one case the relation is particularly clear, viz. in the regulation of water balance. Removal of the posterior lobe of the pituitary causes the condition of *diabetes insipidus*, characterised by the passage of enormous volumes of urine. But lesions in the neighbourhood of the supra-optic nuclei, or in the supra-optico-hypophyseal tracts, also cause polyuria, by interference with the tract connecting this nucleus with the posterior pituitary.† It may be inferred from this that the supra-optic nuclei exert control over the formation, by the posterior pituitary, of the hormone which controls excretion of water by the kidney.

Other functions are also related to the anterior pituitary body; thus injury of the tuber causes sexual dystrophy with adiposity, injury to the anterior region affects carbohydrate metabolism, causing a lowering of blood sugar, etc. Whether such injuries produce these effects by interfering with an innervation which controls the formation of the responsible hormones in the pituitary body is at present uncertain.

The more purely nervous effects of lesions of the hypothalamus are shown by an upset of various intricately co-ordinated functions. One of these is temperature regulation, which will be further discussed in a later chapter. It need only be stated here that hypothalamic lesions, according to their extent and position may cause an animal to lose the ability to withstand heat, cold, or both.

Another complex effect of lesions is the production of a state of somnolence, or, in some species, of sleeplessness. As pointed out in Chapter XXI, the hypothalamus, and especially its posterior and lateral portions, directs the activities of those organs and tissues which are aroused in times of emergency, and it can also facilitate motor responses at various levels, and the withdrawal or inhibition of this drive, as by the inhibitory effect of a "sleep" centre, results in emotional apathy or somnolence.

Lastly, it has been found in man that psychic disturbances are common in patients with lesions of the hypothalamus. These disturbances may take the form of acute mania, or, more rarely, of somnolence and lassitude.

---

* Cushing. "The Pituitary Body, Hypothalamus, etc." 1932. (Thomas, Springfield, Ill.)

† Fisher, Ingram and Ranson. "Diabetes Insipidus and the Neuro-hormona Control of Water Balance." 1938. (Edwards Bros. Ann Arbor.)

For Reference

Beattie (1938). Functional Aspects of the Hypothalamus. Henderson Trust Lectures. (Edinburgh : Oliver and Boyd.)

Collin (1937). L'innervation de la glande pituitaire. (Paris : Hermann & Cie.)

Fulton and Ranson (1940). The Hypothalamus and Central Levels of Autonomic Function. (Baltimore : Williams and Wilkins Co.)

Res. Publ. Assoc. for Res. in Nervous and Mental Dis. (1940). *The Hypothalamus*, **20**, 1–971.

CORTICAL CONTROL OF AUTONOMIC FUNCTIONS

There is abundant evidence that autonomic functions are subject to some degree of control from the cerebral cortex. This is not surprising, since autonomic activity so often accompanies that which is controlled from the cerebrum. In general, the autonomic functions are controlled from areas of the cortex which are concerned with the somatic activities of adjacent structures. Hence, when the motor area for eye movements is stimulated, pupil changes, and lachrymation results, or if that concerned with tongue movements is excited, salivation is seen. Vascular responses, *e.g.* vaso-constriction or vaso-dilatation, changes in heart rate, etc.,* are obtained from certain points on areas 4 and 6, stomach and intestine movements, and sweating, from area 6, etc. These cortical areas appear in part to exert their effects *viâ* the hypothalamus, but partly by other extra-pyramidal tracts.

The Irradiation of Autonomic Reflexes. We have seen that somatic reflexes tend to be directed with some precision, and often to structures closely related to the site of a stimulus. Nevertheless, there is also a tendency for the effects to spread or irradiate to other structures, usually in a definite order.

In the autonomic system, irradiation is a much more pronounced phenomenon. Indeed, as regards the sympathetic system, the effect of an afferent impulse is to set the whole sympathetic system into activity, and its structure is well suited for such widespread responses. Hence, for example, if the central end of the splanchnic nerve is stimulated, the effects reach even to the pupil, which dilates. In the parasympathetic system there is less irradiation than in the sympathetic, but it is still well marked. Thus if the central end of the depressor nerve is stimulated, not only is the heart slowed, but also the peristaltic movements of the intestines are augmented.

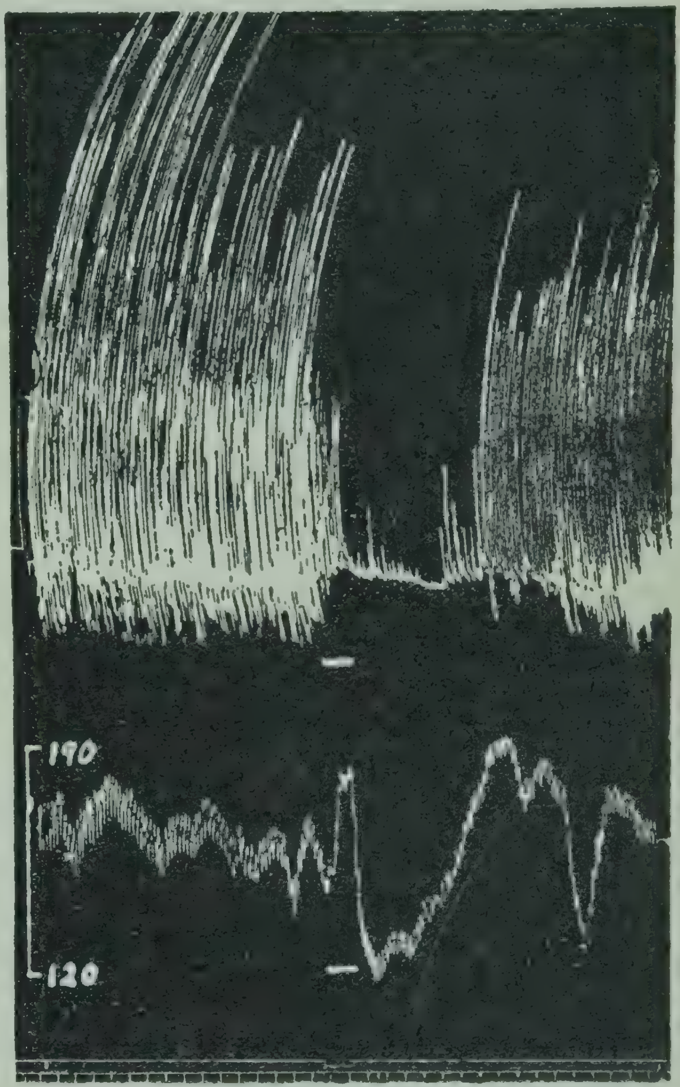


Fig. 238. Chloralosed Cat. Knee-jerks elicited every ten seconds by tapping patellar tendon (upper tracing). Blood pressure in carotid artery (lower tracing). At the signal, stimulation of central end of left vagus. (Schweitzer and Wright. *J. Physiol.*, 1937, **88**, 459.)

Usually the sympathetic and parasympathetic systems act reciprocally, an increase of the one being accompanied by a reduction in the activity of the other. So, for instance, if the nerve from the carotid sinus (Chapter XXXIV) is stimulated, the heart is slowed, partly because the vagus tone is increased and partly because the sympathetic tone is diminished.

---

* Hsu, Hwang and Chu. *Amer. J. Physiol.*, 1942, **137**, 468.

Not only do autonomic reflexes irradiate over the whole autonomic system, but the spread also involves the somatic reflexes. Thus, stimulation of the central end of the vagus nerve causes reduction or abolition of the knee-jerk, and of other somatic reflexes (Fig. 238). There is, however, the possibility that these results may be due to circulatory changes in the centres. Such effects have been shown in the linguo-maxillary reflex (opening of jaw when tongue is stimulated) when the arterial pressure is altered.[*]

For Reference

SCHWEITZER (1937). Die Irradiation autonomer Reflexe. (Basel: Karger.)

THE CHEMICAL TRANSMISSION OF NERVOUS ACTION (THE HUMORAL THEORY)

The similarity between the effects of sympathetic stimulation and those of the administration of adrenaline have long been known, and have given rise to various explanatory speculations. More recently the similarity between the effects of stimulation of the various nerves belonging to the "parasympathetic" system on the one hand, and the action of acetyl choline on the other hand, have excited attention; some instances are given on p. 197 and in Chapter XXXII. As a generalisation bringing these various facts into line, there has been propounded the "humoral" theory, or theory of the chemical transmission of nervous action.

According to this theory, the peripheral effects of sympathetic nerve stimulation are due to the release, *at the nerve endings* concerned, of an adrenaline-like substance, while those of "parasympathetic" stimulation are due to the liberation of acetyl choline at the nerve endings.

The classical instances of sympathetic chemical transmitters were provided by the researches of Cannon and his associates.[†] They found that when the splanchnic or the nerves to the liver were stimulated (the suprarenals having been previously removed), the heart of the animal (previously denervated) was accelerated, as it would have been had adrenaline been injected into the circulation. A similar substance was liberated when the sympathetic nerves to the cat's tail were stimulated, and was believed to have been formed in the smooth muscle of the *arrectores pilorum* muscles and blood vessels.

Denervation of sympathetically innervated tissues (by extirpation of the appropriate ganglion) renders them exceedingly sensitive to the sympathetic transmitter substance, and to adrenaline. A useful test-object is the nictitating membrane of the cat, denervated by removal of the superior cervical ganglion. The suprarenals are also removed. If now any important sympathetic nerve be stimulated, contraction of the nictitating membrane occurs shortly after, showing that a sympatho-excitor substance has been carried by the blood stream.

Once it has entered the blood stream, the sympathetic excitant, like adrenaline, is fairly stable. In the tissues, at its site of liberation, it is liable to destruction by the action of oxidizing enzymes, and hence the action it exerts is fugitive, and more localised than it would otherwise be. The substance is liberated during activity, not only at the endings, but along the whole length of the relevant neurone.

The action of the sympathetic substance, to which the name *sympathin* was given, like that of adrenaline, is augmented by administration of cocaine: possibly there may be more than one substance, and if so, one may be adrenaline itself; another, extractable from spleen, sympathetic nerves and heart, appears to be *nor*-adrenaline, *i.e.* adrenaline deprived of its methyl group.[‡] This probably accounts for the fact that sympathin differs from adrenaline in some of its properties.

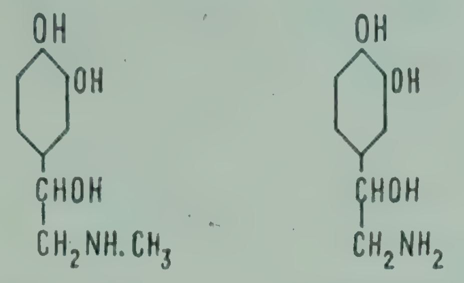
---

[*] ROBILLARD. *Rev. Canadienne de Biol.*, 1924, I, 179.

[†] *Amer. J. Physiol.*, 1931, **96**, 392; **97**, 365; 1932, **99**, 398.

[‡] v. EULER. *J. Physiol.*, 1946, **105**, 38. *Act. Physiol. Scand.*, 1946, **11**, 168; **12**, 73.

Its properties, however, agree precisely with those of *nor*-adrenaline (arterenol).*



Adrenaline *nor*-adrenaline

Nor-adrenaline has an invariably constrictor effect on blood vessels, and this action is annulled, but never reversed, by ergotamine or dibenamine.† Adrenaline has both constrictor and dilator effects, the former usually overpowering the latter: the constrictor effect of adrenaline is therefore reversed by ergotamine or dibenamine, which eliminates the vasoconstrictor component of its action.

"Sympathin" produced by stimulation of the hepatic nerves resembles adrenaline in causing acceleration of the heart, contraction of limb and spleen volumes, and contraction of the nictitating membrane, yet, unlike adrenaline, it does not powerfully dilate the pupil or relax the uterus, and its action on blood pressure is not similarly reversed by the administration of ergotoxine.

In order to attempt to bring these facts into line, Cannon and Rosenblueth ‡ postulated the formation, in the cells affected by the sympathetic, of two forms of sympathin, namely, sympathin E, which is formed when the action is an excitatory one, and sympathin I formed when the effect is inhibitory. It would probably be thought now that adrenaline and *nor*-adrenaline are the substances concerned, the latter corresponding to sympathin E. Whether adrenaline as well as nor-adrenaline is liberated from some sympathetic terminals, other than adrenal medulla cells, is undecided.

The classical instance of humoral transmission of nervous action is that of the liberation of acetyl choline in the heart when the vagus is stimulated § (*v.* Chapter XXXII). The "parasympathetic" transmitter is believed to be identical with acetyl choline. This substance has been found in nervous and other tissues, and in the suprarenal gland; it has also been shown to be present in the blood coming from organs, such as the heart,|| salivary glands, stomach, etc., when their parasympathetic supply is stimulated. After section of the relevant nerve, the endings lose their source of acetyl choline in about 48 hours, and are unable to produce more. Not only is it found in connection with the parasympathetic system, but has also been shown to be released in the cells of sympathetic ganglia ¶ (and also in the medulla of the suprarenal,** which, morphologically, corresponds to a sympathetic ganglion), when these are thrown into action by nerve stimulation, and to be resynthesised in sympathetic ganglia after a discharge.†† Acetyl choline is thus the mediator, not only of the peripheral effects commonly called parasympathetic, but *also of the effects at all pre-ganglionic connections of the autonomic system.*

From the above considerations, it was suggested by Dale that nerves which act by the production of acetyl choline at their terminations, *e.g.* the parasympathetic at their terminals, or the pre-ganglionic autonomic fibres at their synapses in the ganglia or in the suprarenal, should be classed together under the name *cholinergic,* while those that liberate an adrenaline-like substance at their terminals should be called *adrenergic* (Fig. 238). The nerves to the sweat

---

[*] GADDUM and GOODWIN, *J. Physiol.*, 1947, **105**, 357.

[†] FOLKOW *et al. Act. Physiol. Scand.*, 1948, **15**, 412.

[‡] CANNON and ROSENBLUETH. *Amer. J. Physiol.*, 1933, **104**, 557; 1935, **113**, 251.

[§] LOEWI. *Pflüger's Arch.*, 1921, **189**, 239.

[||] FELDBERG and KRAYER. *Arch. F. exp. Path.*, 1933, **172**, 170.

[¶] FELDBERG and GADDUM. *J. Physiol.*, 1934, **81**, 305.

[**] FELDBERG, MINZ and TSUDZIMURA. *J. Physiol.*, 1934, **81**, 286.

[††] KAHLSON and MACINTOSH. *J. Physiol.*, 1939, **96**, 277; FELDBERG. *J. Physiol.*, 1943, **101**, 432.

glands are cholinergic in some animals, such as the cat, but adrenergic in others, *e.g.* the horse, though in all cases anatomically sympathetic and post-ganglionic.

As we have seen in Chapter XII, acetyl choline may also be the humoral transmitter from the motor nerve, *viâ* its end-plates, to skeletal muscle.

The actions of acetyl choline, like the effects of nerve stimulation, quickly pass away, and this is because the substance is rapidly destroyed by enzymes, present in the tissues, and called choline esterases *: in presence of eserine, however, the choline esterase is inhibited, so that the effect of acetyl choline is then more intense and prolonged. There are "*true*" and "*pseudo*" choline esterases, the latter being non-specific esterases, of slow action. True choline esterase alone will hydrolyse acetyl-β-methyl choline, while benzoyl-choline is only split by pseudocholine esterase. There are also specific inhibitors: di-isopropylfluorophosphate (D.F.P.) inhibits "pseudo-", and *pp'*-dichlorodiethyl methylamine hydrochloride (nitrogen mustard) inhibits "true" choline esterase. The inhibition by D.F.P. is irreversible.

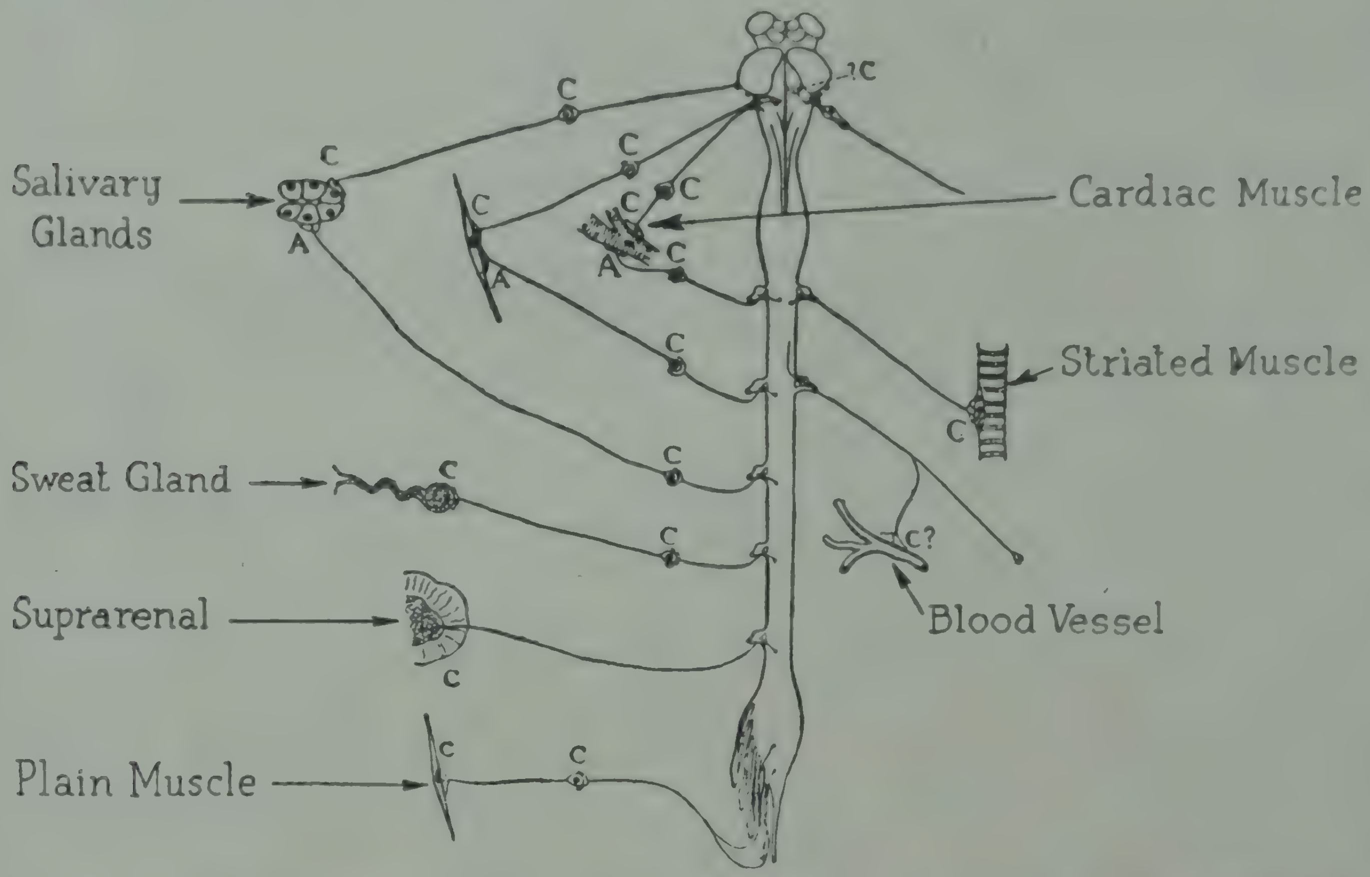


FIG. 239. Diagram of Peripheral Nervous System. At points marked C there is evidence of a cholinergic transmission, at those marked A of an adrenergic transmission. (DALE. *B. M. J.*, 1934, (i.), 835.)

A usual test-object for the detection of acetyl choline in the blood is the dorsal muscle of the leech, previously treated with eserine; this gives powerful contractions in the presence of minute traces of acetyl choline ($1 : 10^9$).

The rapid destruction of acetyl choline at or near to the site of its formation ensures that its action is not only brief, but sharply localised. When its destruction is inhibited by eserine, this brief localised action is lost, and it is then found that stimulation of a parasympathetic nerve, *e.g.* the chorda tympani, not only produces a prolonged effect on the local effector organ (secretion of saliva), but also a generalised effect (fall of arterial blood pressure), owing to the passage of acetyl choline, formed at the stimulated parasympathetic endings, into the general circulation.

Acetyl choline rapidly disappears in alkaline solution. Those of its actions which simulate the action of the parasympathetic nerves are abolished, as the effects of stimulation of those nerves are also abolished, by the action of atropine. The nicotine-like effects, however, *e.g.* the contracture of certain skeletal muscles, and the excitation of the cells of autonomic ganglia, are not paralysed by atropine.

For purposes of physiological identification it is usual to consider the presence of acetyl choline as proved if the given solution causes contraction of the eserinised leech muscle, and has this property destroyed by previous treatment with alkali or with uneserinised blood. Sometimes other tests are used instead, or in addition; these are the contraction of the eserinised rectus abdominis muscle of the frog (by 1 : 10^8), inhibition of the frog's heart (by 1 in 10^8), and fall of arterial pressure in the cat. The last two effects are abolished by atropine which renders the tissue insensitive to A.ch. By comparing these effects with those of solutions of acetyl choline of known strength, a rough quantitative estimation of the substance can be carried out.

---

* GLICK. *J. Gen. Physiol.*, 1938, **21**, 289 (estimation method).

In this way it has been found possible to detect and estimate the acetyl choline liberated in a given organ or tissue by parasympathetic stimulation. All that is needed is to perfuse the tissue in question with eserinised blood or Ringer's solution; the effluent fluid is assayed before and after stimulation of the parasympathetic supply. The use of these methods is illustrated in Fig. 240, which shows that the fluid perfusing a stimulated sympathetic ganglion contains acetyl choline, presumably formed at the synapses.

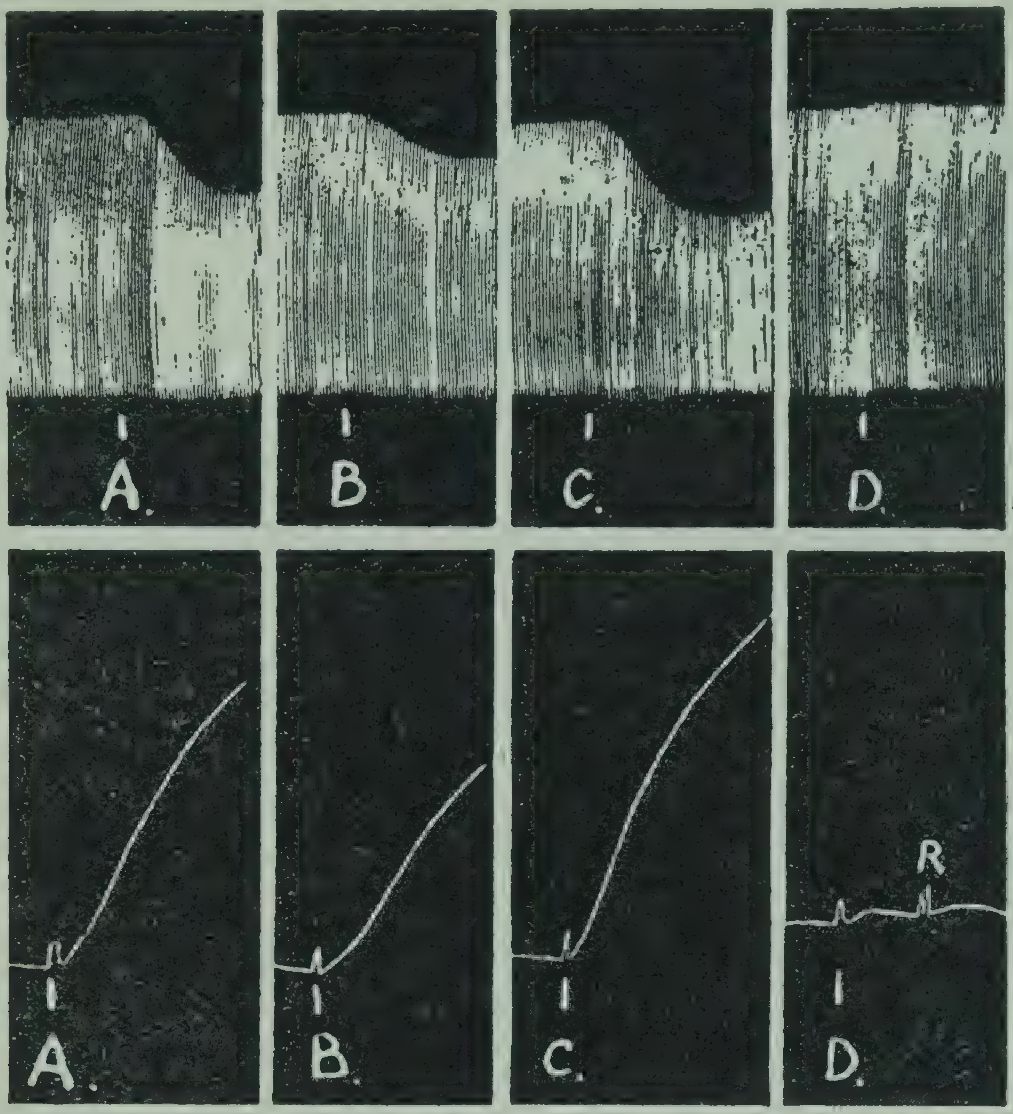


FIG. 240. Demonstration of Presence of Acetyl Choline in perfusate from the superior Cervical Ganglion of the Cat during Stimulation of the Cervical Sympathetic Nerve. Above, effect of perfusate from ganglion on perfused frog's heart; below, effect on eserinised leech. A. Fluid collected during stimulation; B. and C. Comparison by addition of acetyl choline (15γ and 30γ per l = 1 : 15 × 10^9 and 1 : 30 × 10^9 respectively). D = Control fluid without stimulation. (FELDBERG and GADDUM. *J. Physiol.*, 1934, **81**, 305.

It is tempting to suppose that the passage of nervous impulses across synapses in the central nervous system is also effected by the mediation of chemical transmitters. Acetyl choline has been demonstrated in brain perfusates, the liberation being increased by KCl.* The investigations of the Sherrington school on reflex action led to the belief that the phenomena of recruitment, after-discharge and inhibition may be explained on the supposition that some central excitatory state, or some chemical substance causing such state, persists at the synapse for an appreciable interval of time. On the other hand, if such were the case, it would be expected that the arrival of an impulse along a pre-ganglionic fibre might give rise to a prolonged activity in the post-ganglionic neurone, and that only after appreciable delay to allow for diffusion, etc. Yet it has been shown by Eccles that for each pre-ganglionic impulse there is one post-ganglionic impulse, and the delay is very brief at the synapse. Hence, if we are to accept the chemical transmission theory, it is evident that the rate of diffusion and subsequent destruction of the exciting substance must be very great. But in view of the short distances to be traversed and of the extreme rapidity with which such substances as acetyl choline are destroyed by tissues, these objections need not be held to be insuperable.

---

* CHUTE, FELDBERG and SMYTH. *Q. J. Emp. Phys.*, 1940, **30**, 65.

As we have seen, acetyl choline is liberated in a ganglion when the pre-ganglionic fibres to it are stimulated. The effects of intra-arterial injections of acetyl choline and eserine on the C.N.S. have not been shown to be uniform, though in some instances they produce a discharge, inhibited by atropine [*] and facilitated by adrenaline. Acetyl choline, whether injected intra-arterially *viâ* the vertebral artery, or applied to the floor of the IVth ventricle, causes a hyperactivity of the respiratory centre, which resembles that in a normal hyperpnœa and is subject to similar reflex control.[†]

The synthesis of acetyl choline [‡] is brought about by an enzyme "*choline acetylase*," which can be demonstrated in dried nervous tissues, placenta, etc. It requires the presence of K and Mg ions, choline, cysteine, A.T.P., citrate, and an unidentified activator. (NaF is also added to protect A.T.P. from destruction by phosphatase, and eserine to protect the acetyl choline from destruction by choline esterase.)

Acetyl choline is synthesised in the tissues in which it occurs. The choline is probably derived from lecithin; the origin of the acetyl group is unknown, but it may be derived from citrate. As it is formed, the acetyl choline is linked on to some stabiliser in the tissue, probably to a lipoid-protein complex, from which it is released more or less continuously at a slow rate, and suddenly on the arrival of the appropriate stimulus. After release it is rapidly destroyed by the action of choline esterase.

Different parts of the nervous system synthesise acetyl choline in very different amounts, as the following table shows:—

SYNTHESIS OF ACETYL CHOLINE BY NERVOUS TISSUES OF DOG [§]

	Relative amount of A.C.		Relative amount of A.C.
Ventral roots	236	N. gracilis and cuneatus	86
Motor nuclei, X, XII	178	Area 4 (motor)	71
Ventral horns	153	Vestib. nerve	29
Caudate nucl.	127	Int. caps. post. part	27
Lat. geniculate b.	105	Pyramids	12
Dorsal horns	103	Fun. grac. and cun.	10
Thalamus	100	Cerebellar cortex	9
Facial nerve	93	Optic nerve	6

(The amounts are p.c. of the amount formed by the thalamus in the same brain. The actual amount formed by the thalamus was 200–500 μg. per g. dry tissue/hr.)

From this table it will be seen that synthesis is most rapid in the lower motor neurones. The sensory nerves and posterior columns have scarcely any synthetic power, nor have the upper motor neurone (pyramids), the cerebellar cortex, or the afferent parts of the internal capsule. If we suppose that high powers of A.C. synthesis imply cholinergic functioning, and low powers a non-cholinergic mechanism, we are led to the interesting speculation (but it is merely a speculation!) that, in general plan, *successive neurones in the C.N.S. are alternately cholinergic and non-cholinergic*, as illustrated in Fig. 241.

---

[*] BÜLBRING and BURN. *J. Physiol.*, 1941, **100,** 337; CALMA and WRIGHT. *J. Physiol.*, 1944, **103,** 93.

[†] GESELL, HANSEN and WORZNIAK. *Amer. J. Physiol.*, 1943, **138,** 776.

[‡] FELDBERG and MANN. *J. Physiol.*, 1945, **104,** 8.

[§] FELDBERG and VOGT. *J. Physiol.*, 1948, **107,** 372.

The contrast is most striking in nerves or pure fibre tracts, *e.g.* ventral roots and optic nerve. In the grey matter the values are intermediate, because this contains always some cholinergic tissue, either as synaptic endings of one neurone or as the cyton at which another one begins. The nature of the non-cholinergic transmission, whether it is electrical, or adrenergic, or some as yet unknown form of mechanism, must be left open for the present; but the curious effects of adrenaline on synaptic transmission, small amounts tending to potentiate acetyl choline transmission and higher concentrations to depress it; acetyl choline tending to facilitate flexor, and to depress extensor reflexes, while adrenaline has the reverse action *; all these are probably other aspects of a complex picture: there will be little room for surprise, if this alternation speculation should be substantiated, at the conflicting results which have been obtained by the injection of sundry active substances into the general circulation.

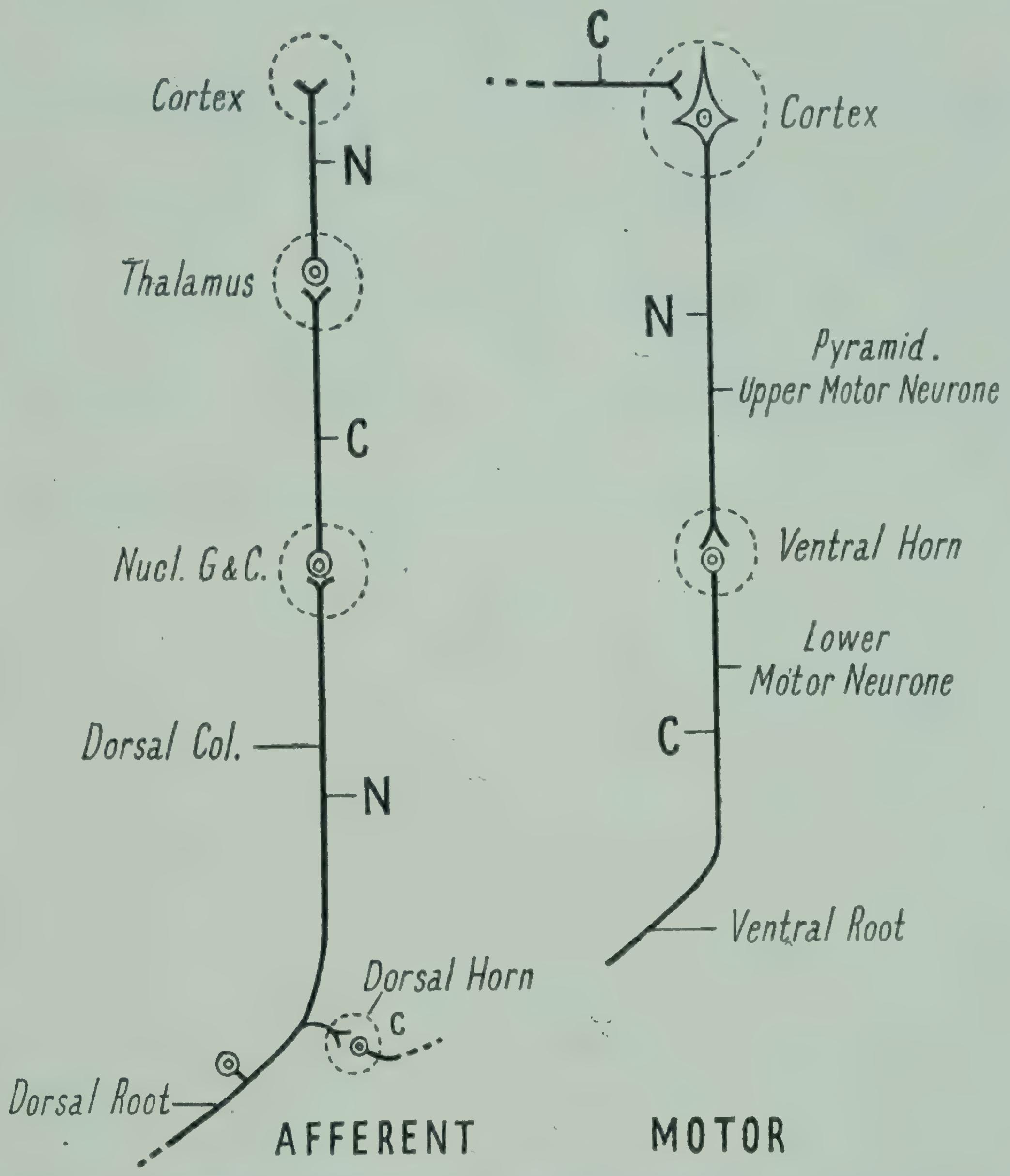


FIG. 241. To illustrate the alternation of cholinergic (C) and non-cholinergic (N) neurones in the central nervous system. Grey matter enclosed by broken lines.

The acetyl choline mechanism has also been supposed to be responsible for the transmission of the nerve impulse along nerve fibres, by causing the depolarisation which essentially constitutes the impulse. Choline esterase is concentrated at the surface of the fibres and is not present in the axoplasm. It is therefore in the best position for removing the A.Ch. very rapidly and so permitting recovery. It may be that this is so for the fibres of cholinergic neurones, but the non-cholinergic nature of the sensory nerves would then mean that the biochemical processes involved in the transmission of the nervous impulse differ fundamentally in the two types.

---

* BÜLBRING, BURN and SKOGLUND. *J. Physiol.*, 1948, **107**, 289.

There are various unexplained phenomena in connection with the theory of chemical transmissions. Thus, atropine abolishes the effects of acetyl choline, and nearly all the effects of parasympathetic excitation. But not all of them – for instance, atropine does not abolish the vasodilatation in the submaxillary gland on stimulation of the chorda tympani, though it does abolish the secretion, yet it abolishes both secretion and vasodilatation on adminitration of acetyl choline.

Ionic changes are certainly also concerned in the phenomenon. Thus the presence of K ions in the blood causes a release of acetyl choline from ganglion cells, but in absence of Ca ions K ions fail to do so ; yet in absence of Ca ions the cells spontaneously discharge.*

For Reference

Brown (1937). Transmission at Nerve Endings by Acetyl Choline. *Physiol. Rev.*, **17**, 485.

Burn (1945). The Relation of Adrenaline to Acetylcholine in the Nervous System. *Physiol. Rev.*, **25**, 377.

Dale (1934). Chemical Transmission of the Effects of Nerve Impulses. *B.M.J.*, (i), 835.

Feldberg (1945). Present Views on the Mode of Action of Acetylcholine in the Central Nervous System. *Physiol. Rev.*, **25**, 596.

Nachmansohn *et al* (1946). The Physico-chemical Mechanism of Nerve Activity. *Ann. N.Y. Acad. Sci.*, **47**, 375.

Rosenblueth (1937). The Transmission of Sympathetic Nerve Impulses. *Physiol. Rev.*, **17**, 514.

---

* Harvey and McIntosh. *J. Physiol.*, 1940, **97**, 408.

BOOK IV

THE SPECIAL SENSES

(BY H. HARTRIDGE, M.D., Sc.D., F.R.S.)

CHAPTER XXIV

THE NERVOUS MECHANISM OF SENSATION

GENERAL PHYSIOLOGY OF SENSATION

THE ability to originate sensations is common to almost all parts of the body. Thus the muscles, joints and viscera send nerve impulses to the brain, which records their activities. We can classify the sense organs into two main groups, those which belong to the common sensibility of the body, and those which form the special senses. The characteristics of the former are (1) that rarely do their impulses, other than those of pain, ever pass the threshold of consciousness; (2) that they play a very large part in the initiation of the numerous reflex actions; (3) that when their impulses do reach consciousness they lack definition. The characteristics of the special senses on the other hand are :—(1) that most of their nerve impulses reach consciousness; (2) that the connection between the stimulus and the resulting response on the part of the individual (as expressed by movements of limbs, etc.) is usually not of the simple nature of a spinal reflex; (3) that the information supplied to consciousness is very definite both as to quality and intensity.

The indefiniteness of the connection between stimulus and response causes us to depend almost entirely on introspection in our study of the sense organs. But introspection as a method of research has grave disadvantages, for we cannot measure our sensations in terms of physical units. We cannot say how red a rose is or how nice it smells. Neither can we compare the intensity of a beam of light, when we feel its warmth on our skin, with the strength of the stimulus which we receive when it enters our eye. Still less can we judge accurately of the sensations evoked in other individuals or animals when we apply stimuli to them. Because of this difficulty of measuring sensations, two methods of investigation have been developed which to a great extent avoid it, namely the *method of threshold values* and the *method of comparisons*.

In the method of threshold values, gradually increasing stimuli are applied to the sense organ under investigation until a sensation is just perceived. Thus, increasing weights are applied to the skin until their pressure is felt, or a light of increasing intensity is presented to the eye until its presence is seen. Such threshold values give definite information concerning the sense organ to which they refer, and allow us to compare the same effects in one and the same sense organ under different conditions and also in different individuals.

In the method of comparisons two separate stimuli are applied, and one or both are varied until the sensations caused by them are equal. Thus a ray of yellow light may be compared with a ray mixed from red and green light, and the intensities of the two adjusted until the sensations caused by both are similar. Both these methods give concordant results if care be taken to make the conditions standard, and both are for this reason largely used for studying the sense organs.

A careful study of these organs is necessary and important because it is only through their agency that we derive information about ourselves, one another, and the world in which we live. The limitations of our sense organs, therefore, restrict our knowledge of the many transformations of energy that are going on around us, except in so far as we are able to devise means of extending them artificially. Thus our knowledge of the existence of ultra-violet light began only with the discovery of photography. Wireless waves circulated through the ether around unknown to us before the invention of the coherer.

Classification of Sense Organs. Sense organs differ in their anatomical position and structure, and also according to the nature of the stimulus to which they react and the kind of sensation which they cause. Any of these differences might be used as a basis of classification, but the latter is found to be the best.

Structure.	Sensations.	
Skin	Touch Pain Heat Cold	Feeling
Eye	Light Colour Shape Distance	Sight
Ear	Pitch Intensity Position	Hearing
Tongue	Acid Sweet Bitter Salt Alkaline	Taste
Nose	Various	Smell

Stimuli adequate for one sense organ are found to be inadequate for another. Thus sound applied to the ear is adequate, but to the eye is not.

The sense organs appear to be elaborated from a very simple type, and the higher they are developmentally, the more is one kind of stimulus adequate and the less are all others.

Thus the end organs in the cornea react alike by a sensation of pain to touch, electricity, heat, inflammation, etc., while the ear reacts to sound waves only. This specific property of the end organs is also attended by an increase in the strength of response to the chosen stimulus. Thus the retina of the eye can be stimulated mechanically and electrically, but has its greatest sensitiveness to light. It is estimated that the eye is many times as sensitive to light as any instrument that has been so far constructed. This specialisation of the end organs makes the information which we obtain from them more detailed and complete, but at the same time sets a limitation to the range of stimulus to which each can respond. Thus a pain end organ can react alike to heat rays, visible rays and ultra-violet rays, while the retina of the eye responds to visible rays only.

The Law of Specific Irritability. A very little consideration suffices to show that there is no resemblance between a sensation and the stimulus, and that one and the same physical event applied to different sense organs will evoke absolutely distinct sensations. If we take a tuning-fork which is vibrating 100 times per second and apply it to the surface of the skin, we get simply a sensation of vibration, *i.e.* a series of tactile impressions repeated at rapid intervals. If the same tuning-fork be applied to the head, its vibrations are imparted to the bones of the skull and thence to the auditory nerve endings and arouse in our consciousness the sensation of a certain musical note. The same thing happens if the vibrations of the tuning-fork are conducted by the ear to the auditory nerve endings in the ordinary way through the external and middle ear. On the other hand, it can be shown that different kinds of stimuli applied to one sense organ always evoke the same kind of sensation, if any. Thus a sensation of light may be aroused not only by the incidence of radiant energy of a certain wave length on the retina, but also by electrical or mechanical stimulation of the retina. If the eye be turned inwards and the finger be pressed on the eye through the outer canthus of the lids, a sensation of light is aroused and we see a coloured circle which we refer to some spot lying to the nasal side of the eye stimulated. The character of the sensation therefore bears no resemblance to the physical events by which the sensation is evoked, but depends entirely on the nature of the sense organ which is stimulated. A sensation of light may be produced by any stimulation of the retina, or of the optic nerve, or of the terminations of the optic nerve in the brain. In the same way stimulation of an auditory nerve or its intracranial endings gives rise to sensations of sound.

Where the question has been investigated, it has not been found possible to evoke different qualities of sensation by different modes of stimulation of nerve fibres. It has therefore been concluded that the quality of any sensation depends simply and solely on the termination of these nerves in the central nervous system, and that where sensations of different quality are produced there must also be difference of nerve paths. This idea was formulated by Müller, and is often alluded to as Müller's "law of specific irritability." The law states that every sensory nerve reacts to one form of stimulus and gives rise to one form of sensation only; though, if under abnormal conditions it be excited by other forms of stimuli, the sensation evoked will still be the usual one.

Localisation. Although the different forms of sensation must be regarded as dependent on the integrity of the brain and of its connections with the peripheral sense organs, sensations are not referred to the brain, but are localised as proceeding from some part of the body or from some region outside the body. Thus the sensation of taste is always localised in the mouth; sensation of touch at the skin or surface of the body; while the sensations of hearing and of sight are "projected," *i.e.* are interpreted as coming from the environment outside ourselves. Even the organic sensations of posture or fatigue are referred to the peripheral reacting parts of the body and not to the central nervous system. A sensation, therefore, cannot be interpreted as a reproduction of external events, but as a *symbol* of these events evoked by stimulation of the sense organs of the body.

Relationship between Stimulus and Sensation. Although the sensation is not a reproduction of the stimulus, it is a symbol of the stimulus, and can be used to inform us of events occurring in the world around. Like stimuli, falling on the same end organ, always evoke like sensations, other conditions being equal. An orderly sequence of sensations may therefore be interpreted as indicating a corresponding orderly sequence of physical occurrences in the world around us. Since our sensations are merely symbols of the physical conditions which give rise to them, it is important to inquire how far they correspond *quantitatively*.

THRESHOLDS. Whatever form of stimulus be applied and whatever sense organs be affected, a certain minimum intensity of stimulus is necessary for it to be effective, *i.e.* to produce a minimum sensation. This strength, which varies with different sense organs, is spoken of as the "liminal intensity" or "threshold value" of stimulus or sensation respectively. In each sense organ we can measure the amount of energy which must be applied to it in order to evoke a minimum sensation. This figure varies considerably with the physiological condition of the subject. Thus it is found to vary in most cases with the following factors:

(1) The physical nature of the stimulus.

(2) The rate of application of the stimulus.

(3) Events occurring in the surrounding areas, *i.e.* simultaneous contrast.

(4) Events which have previously occurred in the area receiving stimulation, *i.e.* successive contrast.

(5) The time during which the stimulus acts, *i.e.* adaptation.

Weber's Law. It is an interesting question how far the strength of sensation may be regarded as an index to the strength of stimulus. Although it is easy to measure in absolute terms the intensity of a stimulus, and although we can say that such and such a light is stronger than another light, it is impossible to say that the sensation resulting from the stronger is two, three or more times that of the weaker. In measuring the effect on sensation of increasing the stimulus, we are therefore reduced to using the *smallest appreciable increase of sensation* as our unit of sensation. Weber's law states that the increase of stimulus, which is necessary to produce an appreciable increase in sensation, always bears the same ratio to the whole stimulus. Thus, if we found that we could just distinguish the difference between a weight of 20 g. and a weight of 21 g., it would not be sufficient to add 5 g. to a weight of 200 g. in order to produce a distinct difference in sensation. In the latter case we should not be able to appreciate any difference until we had added 10 g., *i.e.* one-twentieth of the whole stimulus to the weight.

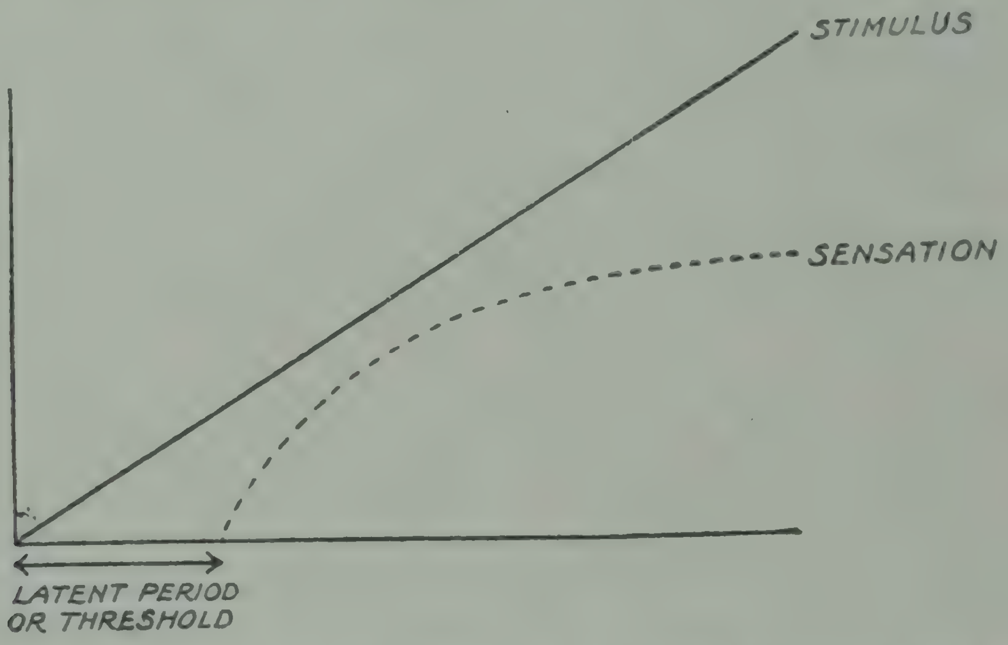


Fig. 242. Diagram to show Relationship between Stimulus and Sensation.

Several methods have been proposed for testing the limits of the applicability of this law. Of these the most important are :

(1) The method of minimal difference.

(2) The method of average error in matching.

In the first method we find by trial how much a given stimulus must be increased in order to evoke an appreciable increase of sensation, and this determination is made for a number of stimuli of different intensities. In the second method it is sought to find the strength of stimulus which is just equal to another stimulus of given intensity. It will be found that errors will be made on both sides, and the average error is taken as representing the minimum difference which is just sufficient to cause a distinct difference of sensation.

In all sense organs Weber's law is applicable only between limits which vary with each sense organ, and it does not hold either for very weak or for very strong stimuli. Within these limits the ratio, which an increase of stimulus must bear to the whole stimulus in order to produce an increase of sensation, may be given approximately as follows for the different sense organs :—

Sense Organ.	*Quality determined.*	*Ratio.*
The ear	Difference of intensity	$\frac{1}{9} - \frac{1}{20}$
The eye	Difference of intensity	$\frac{1}{167}$
The skin	Difference of weight	$\frac{1}{20} - \frac{1}{40}$

Fechner's Law gives the result of an attempt to state Weber's law in mathematical terms. It states that the sensation varies as the natural logarithm of the stimulus. This relationship is shown diagrammatically in Fig. 242.

For Reference

Ass. Res. Nerv. Ment. Dis. (1934). Sensation: its Mechanism and Disturbances (Baltimore: Williams and Wilkins Co.)

CHAPTER XXV

VISION

1. PHYSICAL PROPERTIES OF LIGHT

The Spectrum. It is possible by suitable apparatus to cause the constituent rays in a beam of light to arrange themselves according to their wavelength. When thus arranged they are said to form a spectrum. The apparatus is therefore called a spectroscope. The visible rays thus arranged are seen as a coloured band which has the following appearance. Visibility usually begins at about 8,000–7,800 A.U., the rays of longest wavelength being red. As the wavelength becomes shorter the colour gradually changes to orange, the transition being at about 6,500 A.U. From orange the colour changes to yellow, at about 6,000 A.U.; from yellow to green at 5,500 A.U., to blue-green at 5,000 A.U., to blue at 4,500 A.U., and to violet at 4,000 A.U. The violet extends to 3,800 A.U., where visibility ceases. The spectrum exhibits therefore a gradual change of colour with wavelength. Above the red is the invisible region occupied by the infra-red or heat rays and below the violet the invisible ultra-violet or actinic rays, as explained above.

The colours of the spectrum have important properties which form the foundation of the science of colour mixture. If the spectrum produced from white light is caused to fold up again, it is found that white light is re-formed. But white light is produced if certain pairs of colours only are caused to combine in the correct proportion. Thus red (6,562 A.U.) and blue-green (4,921 A.U.) when mixed correctly form white light, so also do yellow (5,636 A.U.) and violet (4,330 A.U.). Such pairs are called *complementary colours*. But since there is in the spectrum a gradual transition from one colour to the next, so there are between red and yellow an infinite number of rays of different wavelength, each of which has its complementary colours, between blue-green and violet. If, therefore, from white light we remove one of a pair of complementary colours, the other member of the pair will be left unneutralised, and the light thus becomes tinted with its colour. Green rays do not possess a complementary in the spectrum; but it is found by experiment that, by combining red and violet to form purple, the required colour may be produced. If we include purple with the spectral colours, we can imagine these colours to form a closed ring. Each colour will then have its complementary opposite to it.

The Spectral Colours have another important property, for if red and yellow are caused to combine, they are found to produce orange, the intermediate colour. If red and green are mixed, then again the intermediate colour, yellow, may be produced. It is found that by varying the intensities of the two components, it is possible to produce orange or yellow-green, or in fact any other intermediate colour at will. Careful experiment shows that the intermediate colour thus formed is no mere approximation but an exact match. If red and green are thus able to combine to form the intermediate colour, while red and blue-green are complementaries producing white by their mixture, the question arises as to the effect produced by mixing red with a colour intermediate between green and blue-green. Experiment shows that a range of colours is produced containing an amount of white light, which varies with the intensities and wavelengths of the combining colours. Colours diluted with white light are spoken of as *unsaturated*. In order that the colours produced by a mixture of red and green rays shall be fully saturated, and thus be able to match the colours of the spectrum exactly, the green must not be shorter in wavelength than 5,400 A.U. Similar phenomena are to be found at the other end of the spectrum; green and violet, when mixed in various proportions, form colours which match the intermediate spectral colours. With red, green and violet it is therefore possible to match the whole spectrum. But since red and violet, when mixed, form the intermediate purples, it is possible with the three coloured rays to imitate the whole range of colours. Now the purple formed from red and violet is, as we have seen, the complementary colour to green; by means of these three colours it is thus possible to produce white light. It should therefore be possible to match an unsaturated colour as easily as a saturated one. Experiment shows that such is the case. The third property, beside colour and saturation, is *intensity*, which depends on the amplitude of the waves. The intensity of the mixtur formed by red, green and violet, can therefore be readily adjusted by varying the intensit of each of the three component rays. We may summarise the above facts by stating that by varying the intensities of the red, green and violet rays, it is possible to match ever shade and colour. This statement has been put to the test by Maxwell, Abney and othe observers, and has been found to hold good in all cases but one, spectral blue being slightl more saturated than the mixture of green and violet. In describing the complementar pairs of colours, we have mentioned that if the spectral colours are placed in a close ring, complementary pairs are found to be opposite to one another. If now the thre fundamental colours are placed at equal intervals round the ring, we may regard whit as occupying the centre, because it is equidistant from the three fundamentals, and at th same time lies on the diameter between the various colours and their complementaries. If the other spectral colours are arranged in position relatively to the three fundamentals, they form a figure that in shape resembles a triangle more closely than it does a ring. This is due to the facts already mentioned (1) with regard to the exact matching of the spectral colours between red and green, by mixtures of those two fundamental colours; (2) with regard to the approximate matching of the region between green and violet, by the mixtures of those colours; and (3) with regard to the exact matching of mauves and purples, by mixtures of red and violet. The colour triangle which is shown in Fig. 243 therefore has a purely experimental basis, and has no association essential with theories of vision.

Photo-chemical Change, which is the third property of light mentioned above, is well illustrated by photography. The most important principle of light action is that light, to cause chemical change, must be absorbed (Draper's law). For example, an ordinary photographic plate which is found to be opaque to blue, violet and ultra-violet rays, and to be transparent to the rest of the visible spectrum, is therefore sensitive to the former rays but inactive to the latter. Further, by colouring the plate by a dye which absorbs red, yellow and green, it is possible to make the plate react to these rays. Draper's law is therefore obeyed. Chemical reactions caused by light are of many types, but may be divided into reversible and irreversible. The former type of reaction occurs only so long as the light acts (the change from CO- to oxy-hæmoglobin may be given as an example), while the latter type remains in the final state that has been reached (the changes in a photographic plate may be given as an example). There is, further, another and more complicated type which, when once started by an incident beam of light, goes on automatically with an evolution of energy until the reaction is completed. These effects of light are probably of great importance in connection with vision, and will therefore receive further consideration later.

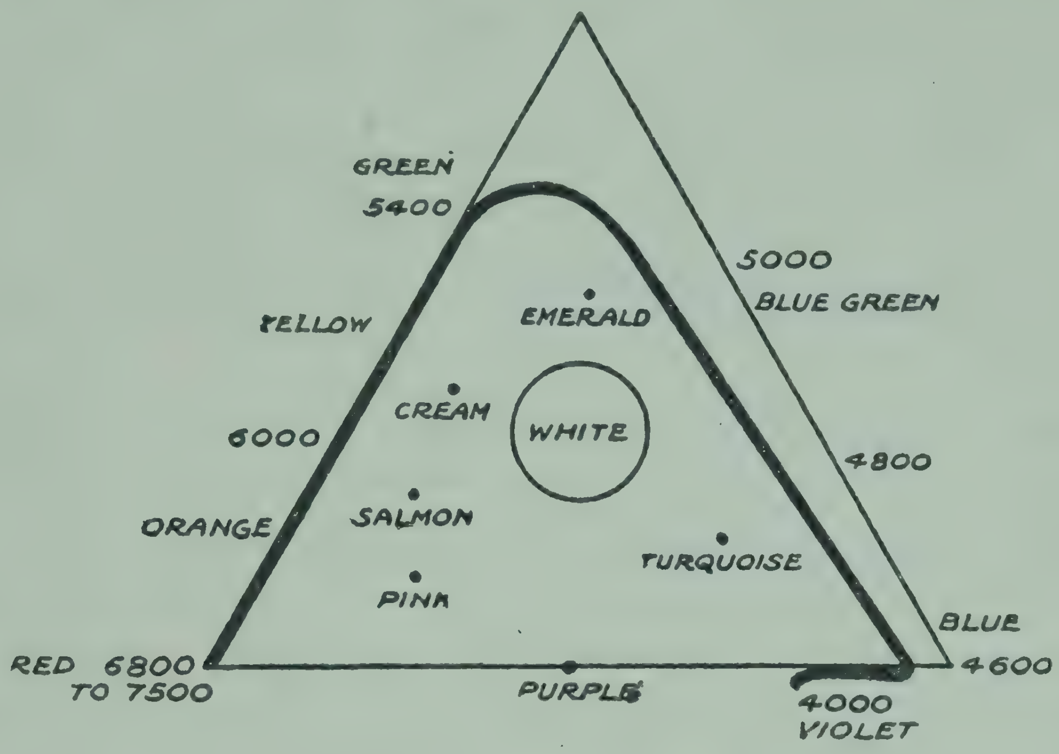


Fig. 243. Colour Triangle.

The black line shows the shape of the curve along which the different rays of the spectrum fall for white to occupy the central position.

2. EYE MOVEMENTS

ANATOMY AND FUNCTION OF THE EXTRINSIC MUSCLES OF THE EYEBALL

Six external muscles produce rotation of the eyeball:—four recti and two oblique. The recti arise from a fibrous ring attached to the margin of the optic foramen. From the positions they occupy they are called superior, inferior, external and internal. When they contract they cause upward, downward, outward and inward rotation of the eyeball respectively. In the case of the first two muscles there is a turning movement inwards at the same time. Fig. 244 shows the above diagrammatically.

The two oblique muscles, the superior and inferior, are both smaller than the recti. The former arises near the optic foramen, and passes forward to the upper and inner side of the orbit, forming on its way a round tendon. It here passes through a narrow fibrous ring, and then turns downwards and backwards under the superior rectus and becomes attached to the eyeball. The inferior oblique arises from the nasal side of the orbit, just within its lower margin. It passes outwards and backwards beneath the inferior rectus to become attached to the eyeball nearly opposite to the attachment of the superior oblique. On contraction of the superior oblique the upper side of the eye is rotated towards the nose; at the same time the pupil is directed downwards and outwards. The inferior oblique also causes rotation about the visual axis, but in the opposite direction; it at the same time produces upward and outward movement of the pupil.

On tilting the head suddenly about a transverse axis, it is found that the eyes rotate in the opposite direction, so that in fact the image formed on the retina shall still keep in the same meridian with respect to external objects. This rotation is called "compensatory" and is largely effected by the oblique muscles.

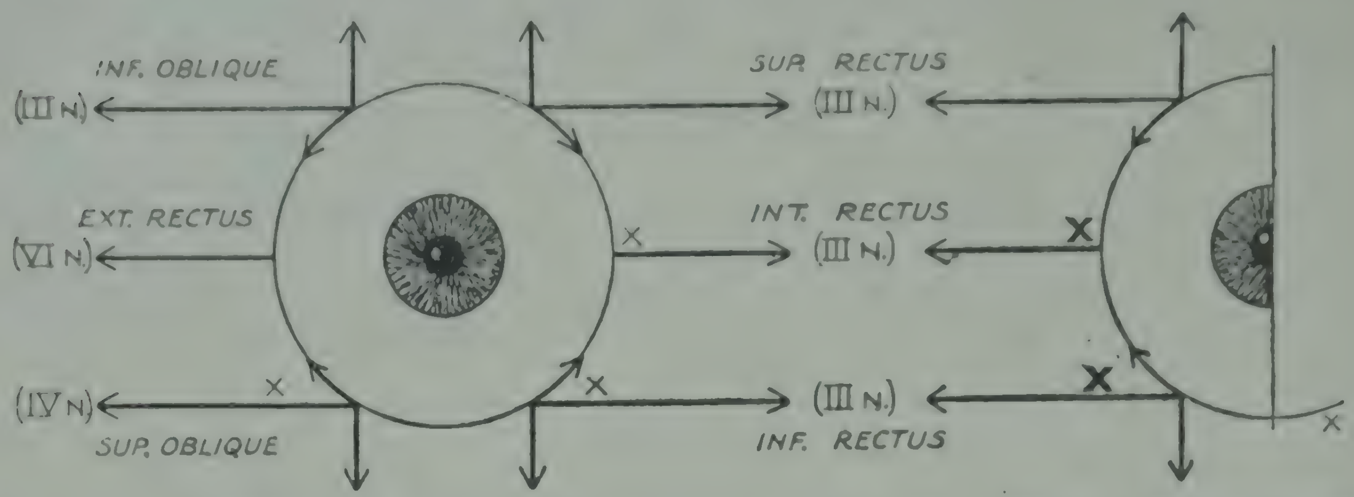


FIG. 244. Diagram showing the Directions in which the different Extrinsic Eye Muscles rotate the Eyeball. x = *crossed connection.*

Co-ordinated Movements of the Eyes. The notable feature of the eye movements is the close association which exists between the muscles of the two eyes. So perfectly has this mechanism been developed that the eyes are able to glance rapidly from place to place without there being any obvious doubling of the images. The eye movements are therefore of such a kind that the image of an object conveys a single impression to consciousness. But objects vary in their distance from the observer, and therefore, beside movements of the eyes in which the visual axes remain parallel, there are also movements in which there is a certain amount of convergence. In the latter case there is usually some associated accommodation of the lens for near objects, and at the same time some contraction of the pupil. By experiments in which prisms are placed in front of the eyes, thus calling for convergence or divergence without accommodation, and by others in which lenses are placed there instead, thus requiring accommodation without change in the angle between the axes, it can readily be shown that the association between the functions of accommodation and convergence is not very rigid. The co-ordinated deviations of the eyes appear to be much more closely connected. Thus Donders found co-ordinated deviations both in the newly born and in congenitally blind persons. This is probably due to the close anatomical relationship which exists between the nerve centres of the muscles on the two sides.

The Control of Eye Movements. The efficiency of the eyes as organs of vision depends very largely on the speed and precision with which their axes can be turned in any required direction. These movements are of three kinds: (*a*) Those in which the eye axes move together to right or left, up or down. (*b*) Those in which the eye axes move in opposite ways, the left eye axis to the right and the right eye axis to the left, thus bringing about "convergence," and *vice versâ* for "divergence." (*c*) Those in which the eyes rotate clockwise or anti-clockwise about their axes. All eye movements can be analysed into combinations of these three components. The movements may be brought into operation in two ways: voluntarily, *e.g.* the directing of the gaze from line to line of print, as in reading; and involuntarily, as when the eyes turn instinctively to scrutinise an object which has presented itself to the field of view, or when the eyes follow an object as it moves; this movement of the eyes is called the "fixation reflex."

The eye muscles are innervated by the third, fourth, and sixth cranial nerves, which are derived from nerve cells found in the grey matter of the mid-brain which lies just ventral to the central canal (or Sylvian aqueduct), at the level of the superior corpora quadrigemina. Fig. 245 shows approximately the positions of the various collections of nerve cells which innervate the different muscles. It will be seen that the third or oculo-motor nerve supplies all the extrinsic eye muscles except three, viz.: Tenon's capsule, which is supplied by the sympathetic; the superior oblique, which is supplied by the fourth nerve; and the external rectus, which is supplied by the sixth nerve. Further, while most of its nuclei supply muscles on the same side, three are found to go to muscles on the opposite side, namely, the internal and inferior recti (third nerve) and the superior oblique (fourth nerve).

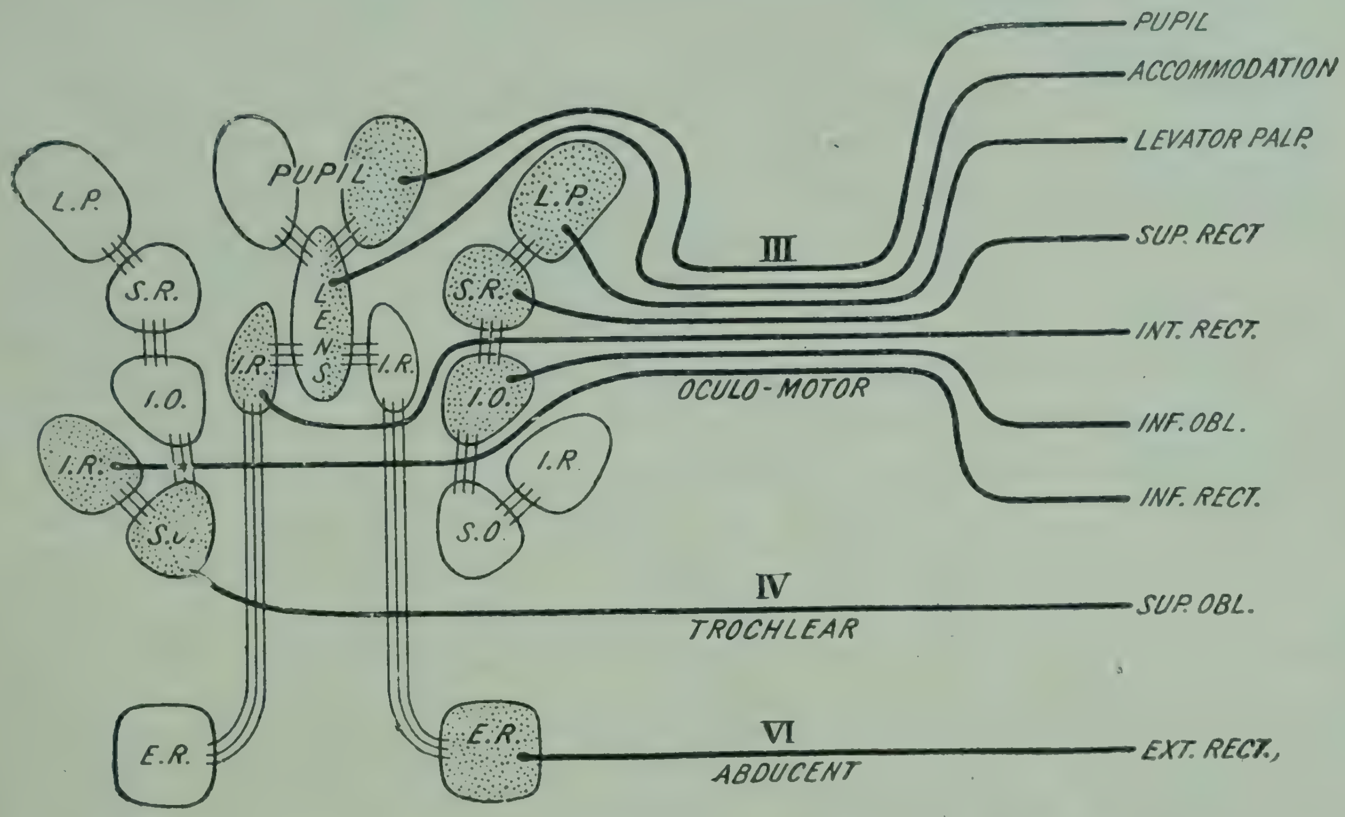


FIG. 245. **Diagram to show Relationships of different Points of Oculo-motor Nuclei, and the principal connections between them.**

Sherrington showed that, as the muscle on one side of a limb contracts, its antagonist at the same time relaxes, so as to allow the movement to take place smoothly and without waste of energy. This is called "reciprocal innervation." The eye muscles show the phenomenon very well. If the right frontal cortex be stimulated, the eyes perform co-ordinate deviation to the left. If now all the muscles of the right eye except the external rectus are divided, it is found that on repeating the stimulation this eye still moves in co-ordination as far as the middle line, through the relaxation of the external rectus muscle.

The orbicularis palpebrarum is also supplied from the third nerve nucleus, for in lesions of this nucleus paralysis of this muscle is found. The fibres innervating it probably travel all the way with the seventh nerve.

CAUSES AND DIAGNOSIS OF STRABISMUS

Squint or strabismus may be caused by a number of conditions: (1) by congenital abnormality; (2) by interference with the proper rotation of the eyeball; (3) by injury to one of the external eye muscles; (4) by injury to or stimulation of one of the nerves supplying these muscles; (5) by the presence of certain errors of refraction With regard to nerve injury the following description may be given. Injury to the third nerve causes (*a*) drooping of the upper lid owing to paralysis of the levator palpebræ; (*b*) external strabismus from paralysis of the upper, inner and lower recti and the unopposed action of the external rectus; (*c*) rotation of the eye about its visual axis from paralysis of the inferior oblique and therefore unopposed action of the superior; (*d*) dilatation of the pupil from paralysis of its sphincter and the unopposed action of the dilator fibres which are innervated by the sympathetic; (*e*) loss of the power of accommodation from paralysis of the ciliary muscle; (*f*) exophthalmos or protrusion of the eye, caused by the paralysis of so many of its muscles and the unopposed action of the smooth muscle fibres in Tenon's capsule. Owing to the fact that for a considerable portion of its course the third nerve lies beside the fourth, fifth and sixth nerves, there is usually also some associated symptom of paralysis in the structures which these nerves supply. Injury to the fourth nerve causes paralysis of the superior oblique, which shows itself by defective movements in a downward and outward direction. Injury to the sixth nerve causes internal strabismus owing to paralysis of the external rectus. This frequently occurs when tumours, hæmorrhage, inflammation or injuries involve the base of the brain. Experimental stimulation of these nerves causes contraction of the muscles supplied, and does not require specific description.

DIPLOPIA is a characteristic feature of a strabismus. Objects appear double to the patient unless or until he has acquired the ability to suppress the images presented to one (usually the weaker) eye. This suppression gives relief from the diplopia, but involve the risk of the loss of function of the unused (suppressed) eye. An eye that has suffered in this way is said to be amblyopic.

Diplopia is also observed when an individual is under the influence of certain drugs, *e.g.* alcohol. It is also seen sometimes when the spherical curvature of the cornea is disturbed by injury or disease, *e.g.* corneal ulcer. In these cases the injured eye sees a double image even when the sound eye is closed ; in fact it may see more than two, the condition being called polyopia.

Voluntary Eye Movements. The impulses for these originate in the Betz cells of that part of the voluntary motor area which has its centre about two-thirds down, and somewhat in front of the precentral (Rolandic) strip of cortex. The axons from these cells pass downwards and inwards to form part of the angle (or genu = knee) of the internal capsule. The fibres now descend vertically with the rest of the pyramidal fibres till they reach the level of the superior colliculi where, crossing to the opposite side (as all the other pyramidal Betz cell fibres cross lower down) they end in relation to the third, fourth, and sixth nerve nuclei, either directly or by means of short intervening relay fibres. By this crossing of the fibres the cortex becomes connected with the nerve nuclei of the opposite side. Now the right eye is turned towards the left by fibres which originate in the left third nerve nucleus (see Fig. 245). The left eye is turned towards the left by fibres which originate in the left sixth nerve nucleus. Both these nuclei are controlled by impulses which have originated in the right cerebral cortex. In consequence, impulses from the right cortex turn both eyes to the left. But the crossing of the other voluntary motor fibres causes the cortex on the right side to control the left side of the body, *e.g.* the left arm and hand. The turning of the eyes to the left is therefore very appropriate, since it will cause the gaze to be turned towards the side on which other activities may be in progress.

The Fixation Reflex. Fixation is the retention of the gaze of the two eyes on an external object. It is effected by such adjustments of the external eye muscles that an image of the external object remains on the foveæ of the two eyes. These adjustments are carried out with extreme precision and dispatch. The reflex develops at a very early age, since babies a few days old perform fixation perfectly. Two sets of paths are involved in the reflex : (*a*) Those from retinæ to calcarine cortex *via* the optic radiation ; (*b*) those from the calcarine cortex to the superior corpora quadrigemina (colliculi), probably on the opposite side, and from them to the third, fourth and sixth nerve nuclei.

The mode of action of this reflex is probably as follows : Suppose an object, *e.g.* an electric lamp, to be voluntarily fixated, the centres of its images will fall on the foveæ of both eyes and the stimulus produced by it will fall partly on the right and partly on the left halves of both retinæ. Both right and left calcarine cortices will therefore receive impulses. From these fresh impulses will travel to both right and left sets of third, fourth and sixth nerve nuclei, probably *via* the superior corpora quadrigemina of both sides. The eyes will not be moved since the nerve nuclei on one side would produce a movement which is opposed by that of the nerve nuclei on the other side. Suppose now that the lamp is moved a little to the right. The images produced by the eye media on the retinæ will move to the left and the left-hand halves of the retinæ of both eyes will, as a result, be more strongly stimulated than the right-hand halves. In consequence, the impulses reaching the left-hand calcarine cortex will be more powerful than those reaching the right. Stronger impulses will therefore pass down to the superior corpora quadrigemina (colliculi) and so to the third, fourth and sixth nerve nuclei of the opposite side, that is of the right side. In consequence, the eyes will be turned towards the right, and this will cause the centres of the images of the lamp to fall once more on the foveæ. In some such way as this we may picture fixation being preserved.

Ordinarily fixation is under voluntary control. When this control becomes defective, fixation is still effected, but the gaze "snaps on to" an external object and remains fixed to it. With a great effort the gaze is turned away from this object only to snap on to another. The use of the eyes, *e.g.* in reading, thus becomes impossible, since the gaze cannot be steered along the line of type.

VOLUNTARY AND INVOLUNTARY EYE MOVEMENTS

Many years ago Hering made the suggestion that visual acuity for certain kinds of test-object would be improved if the eyes were performing tiny rotations during the process of fixation. Since that time the idea appears to have been gaining ground that even normal subjects continually suffer from what may be called "a rapid nystagmus."

This matter has recently been investigated by Lord and Wright, by Barlow, and by Hartridge and Thomson. Lord and Wright record head movements as well as eye movements; so does Barlow. Hartridge and Thomson employ a method which records eye movements independently of head movements. Agreement between these observers is not complete however. The present trend of opinion is that the eyes do not perform this rapid nystagmus which Hering suggested, but that they maintain a fixed position, to a remarkable degree, when careful observations are being performed on a test-object.

MINUTE ANATOMY OF THE EYE. The Cornea has a radius of nearly 8 mm. and a diameter of 11 mm. It is composed of the following five layers: (1) Stratified epithelium. (2) The anterior elastic lamina of Bowman. (3) Substantia propria. (4) The posterior elastic lamina of Descemet. (5) A layer of endothelium.

The Iris consists of three layers: (1) the endothelium, continuous with that on the posterior surface of the cornea; (2) the stroma of the iris, which consists of connective tissue (especially elastic fibres), two thin sheets of muscle, some pigment cells, vessels and nerves; (3) the pigmented layer continuous with the retina.

It should be noted that the posterior elastic lamina of Descemet in the cornea, after its division into three parts, forms by its posterior portion the ligamentum pectinatum iridis, by which the iris gains attachment to the sclero-corneal junction.

The Functions of the Iris. The iris contains two layers of unstriated muscle fibres; the anterior, which is circularly arranged, so that by its contraction it acts as a sphincter; while the posterior is arranged radially, stretching from the attachment of the iris to the rim of the pupil, so that by its contraction it causes the pupil to open. Because of the numerous pigment cells which it contains the iris is opaque to light. Contraction of the pupil thus causes the following effects: (1) reduction in the amount of light entering the eye, so that an image of less intensity is formed on the retina; (2) the use of the more central zones of the lens system only. The advantage of this lies in the fact that, as will be described later, the more peripheral zones suffer from errors of refraction to a much greater degree than do the central ones: the contraction of the pupil therefore improves the definition of the image; (3) an increase in the depth of focus of the eye, which is of great value for near vision. The way that depth of focus is obtained will be described later (see p. 393).

Contraction of the pupil occurs under the following circumstances:

(1) If the third cranial nerve or its nucleus is stimulated electrically, or if the sympathetic nerve fibres to the eyes are cut or are paralysed.

(2) When light falls on the retina. This movement, which is known as the "light reflex," is determined by a contraction of the sphincter pupillæ, together with a relaxation of the dilator muscle. The contraction ensues within a period of 0·04 to 0·05 sec. after the moment at which the light has access to the retina, and attains its maximum within 0·1 sec. In man as well as in other animals which have binocular vision, and in which there is a partial decussation of the fibres of the optic nerves in the optic chiasma, the reflex is bilateral, *i.e.* light falling into one eye causes simultaneous contraction of both pupils. In the higher animals this reaction of the pupil to light demands the integrity of the nervous paths between the eye and the brain; but in many of the lower animals, *e.g.* in the frog and eel, the reflex nervous mechanism is aided by a local sensibility of the iris to light. In

these animals the contraction of the pupil in response to illumination takes place even in the excised eye, and seems to be determined by a direct stimulation of the pigmented contractile fibres of the sphincter pupillæ by means of the light.

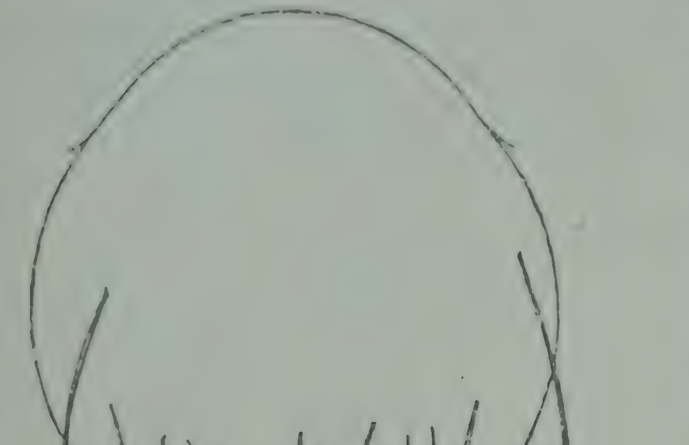


Fig. 246. Nerve Supply to the Eyeball. (After Foster.)

l.g.	Lenticular ganglion with its three roots, viz. :
r.b.	Radix brevis or short root.
r.l.	Radix longus or long root.
sym.	Sympathetic root.
V. opth.	Ophthalmic division of *V.* nerve.
III oc.m.	Oculo-motor nerve.
II.	Optic nerve.
l.c.	Long ciliary nerves.
s.c.	Short ciliary nerves.

The effect of light on the pupil varies considerably according to the condition of adaptation of the eye. The diameter of the pupil is maximal when the eye has been in the dark for some time and may amount then to 7·3 to 8 mm. In one experiment, on exposing the eye to a feeble light, *e.g.* 1·6 candles at a moderate distance, the pupil diminished in size to 6·3 mm. ; with an illumination of 50 to 100 candles the size of the pupil was 3·7 mm., and with 500 to 1,000 candles, 3·3 mm. This effect was obtained by a *rapid* change in the illumination of the eye. When the alteration is sudden, the amount of contraction was found by Haycraft to be equal to the logarithm of the intensity of the light. When, however, the alteration is so gradual that the retina can become adapted to the change as it proceeds, then little or no change in the size of the pupil occurs ; and when the illumination, which has at first caused a maximal constriction of the pupil, is continued, the pupil gradually relaxes with the adaptation of the retina to light. This relaxation occurs within three or four minutes after exposure to light has taken place. The same influence of adaptation will be observed if two individuals are brought into a moderately lighted room, one from bright daylight and the other from a dark room. The pupils of the first will dilate widely, while those of the second will constrict. In each case the change will pass off rapidly, so that at the end of five or ten minutes there will be no difference observable between the eyes of the two persons.

(3) When vision is directed to a near object, the processes of accommodation of the lens, and of convergence of the visual axes, which result, are associated with contraction of the pupil. The sharpness of vision is thereby improved, together with an increase in the depth of focus, a result very beneficial for the close examination of detail. Since it is possible by experiment to cause accommodation without convergence and *vice versâ*, we may ascertain which function is the more closely associated with the pupil mechanism. The evidence appears to be in favour of convergence.

(4) In sleep the pupils are always contracted. This behaviour may enable us to distinguish feigned from real sleep. This contraction of the pupils, in spite of the fact that no light is entering the eyes, has been held to be caused by association with the upward and inward direction of the eye axes which was said to be found in sleep. There now appears to be irrefutable evidence that the eyes during sleep may occupy any position ; another explanation of the constricted pupils must therefore be found.

(5) Contraction of the pupils is a marked effect of certain drugs such as morphine, which acts by stimulation of the third nerve nucleus; other drugs, *e.g.* pilocarpine, muscarine and eserine, facilitate the action of acetyl choline formed at the 3rd nerve endings in the sphincter pupillæ muscle.

(6) Constricted pupils are also met with in excitable conditions of the central nervous system, and therefore during the induction of chloroform and ether anæsthesia.

(7) Small pupils, which do not react to light, are also met with in injuries to the spinal cord involving the cervical region and damaging the pathway of the sympathetic dilator fibres of the iris.

(8) Contracted pupils are found to accompany severe pain. This is probably due to the powerful flow of efferent impulses which leave the brain in this condition, affecting the 3rd nerve nucleus which controls the pupil.

(9) The pupil contracts when the aqueous is allowed to escape from the anterior chamber. The cause of this is said to be the dilatation of the vessels of the iris, owing to the fall of the surrounding pressure.

Dilatation of the Pupil occurs under the following circumstances:

(1) If the sympathetic nerve fibres to the eye are stimulated, or if the third cranial nerve fibres are cut.

(2) On removal of a light stimulus from the eyes. If the removal be complete the pupil remains dilated, but if there be any light at all the pupil gradually contracts again as the eye becomes dark adapted.

(3) On accommodation for distant vision because the associated reflex stimulation of the pupillo-motor centre with accommodation is no longer called into play.

(4) Reflex dilatation of the pupil can be excited by the stimulation of any sensory nerve. This may be due to some of the afferent impulses reaching the cilio-spinal sympathetic nerve centre in the cord.

(5) The pupils are frequently found to dilate in such emotional states as fear, anxiety, exhaustion and dyspnœa, and also at the moment of death.

(6) Dilatation is also found to accompany extreme exhaustion of the central nervous system, when the activity of all nerve centres is low, such as in deep chloroform anæsthesia, and in the coma produced by alcohol poisoning. Many drugs such as atropine and homatropine, cause dilatation by paralysing the endings of the 3rd nerve. Others, such as cocaine and adrenaline, dilate by stimulating the endings of the sympathetic nerves.

(7) Dilated pupils inactive to light are found in injuries of the 3rd nerve or its nucleus.

(8) Dilated pupils are also found when the intraocular pressure is abnormally high, as in glaucoma. This appears to be due to constriction of the vessels of the iris owing to the high external pressure to which they are subjected.

(9) Dilated pupils inactive to light are found in compression and severe concussion of the brain. This is probably due to the abolition of the normal nervous impulses to the muscles, so that the pupil dilates under the influence of its radial elastic fibres.

(10) Dilated pupils are found to accompany hyperactivity of the suprarenal glands, owing to the presence of considerable amounts of adrenaline in the blood. This occurs for example in fear and in oxygen want, dilated pupils being one of the characteristic signs of those conditions.

(11) The pupils are found to dilate reflexly on stimulating the central end of the splanchnic nerve (McSwiney). The impulses enter the spinal cord by the posterior roots as high as the 3rd thoracic segment. They reach the 3rd nucleus and cause inhibition there.

3. THE NOURISHMENT AND PROTECTION OF THE EYE

Anatomy of the Lids. Closing the orbit in front, and in close relationship to the eyes, are the lids or palpebræ. The upper, which is the larger and the more movable, is provided with a special muscle, the levator palpebræ superioris. This is supplied by a branch of the oculo-motor (3rd) nerve. The two lids meet at an angle on both sides, forming the inner and outer canthi. They are stiffened by two plates of dense fibrous tissue, parallel to their edges, which are called tarsi. Near these and embedded in the substance of the lids are two sets of glands, the Meibomian glands and those of Moll. These secrete a greasy material which spreads over the lids. Superficial to these structures, but under the skin, is a ring of muscle fibres which is common to both lids, the orbicularis palpebrarum, innervated by the 7th nerve. Its contraction closes the lids. Lining the inner surfaces of the eyelids is a thin layer of mucous membrane, the conjunctiva, which is reflected on to the front of the eye, and is continuous over the cornea as the anterior epithelial layer.

CLOSURE OF THE LIDS occurs: (1) during sleep; (2) if a very bright light enters the eyes; (3) on the sudden approach of some foreign body; (4) on contact of a foreign body with the lashes; (5) on irritation of the cornea or conjunctiva, *e.g.* by touching (the *conjunctival reflex*); (6) in sneezing; (7) in order to renew the fluid film on the cornea and conjunctiva. The reflex closure of the lids is therefore a very important function in affording protection to the eyes. The reflex apparently can be initiated by the stimulation of any of the branches of the ophthalmic (1st) division of the 5th (trigeminal) nerve. From the nucleus of this nerve in the pons Varolii, fresh fibres take the impulses, it is believed, to the 3rd nerve nuclei of both sides. The fibres leave the 3rd nerve, however, and join the 7th (facial) nerve, from which they pass to the orbicularis palpebrarum. The conjunctival reflex is one of the last to be abolished by anæsthetics and is therefore used as a convenient test.

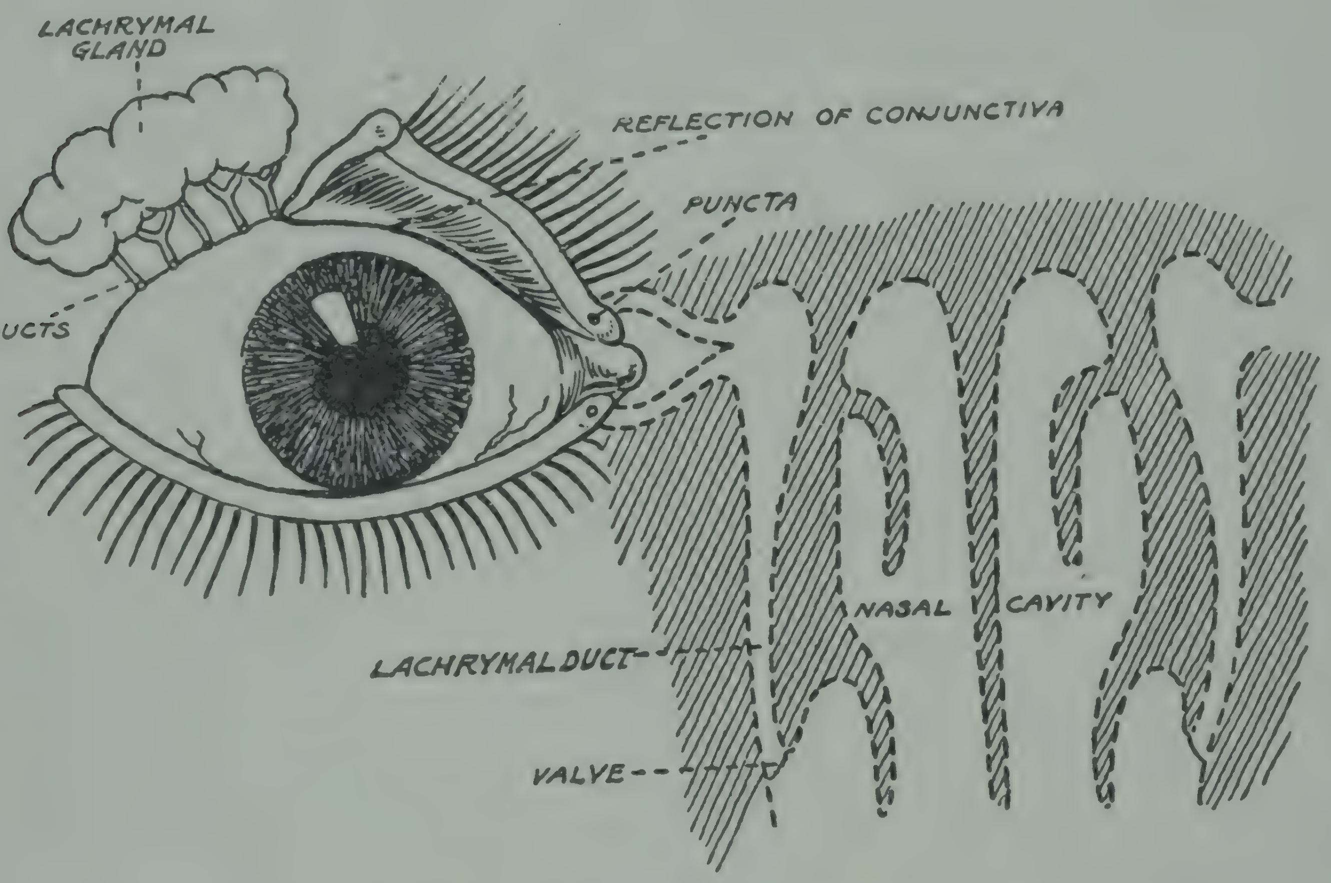


FIG. 247. Diagram to show Origin and Fate of Tear Fluid.

The conjunctivæ and the cornea are kept in a moist condition by the tears, which are secreted by the lachrymal gland, situated in the upper and outer part of the orbit. This is a small acino-tubular gland, in microscopic structure similar to the parotid. Its secretion issues through several ducts, the mucous linings of which are continuous with that of the conjunctiva. Normally the secretion is just sufficient to keep the surfaces of the lids and cornea moist, the production keeping pace with the evaporation. Under certain circumstances there is excess, and tears are produced.

TEAR FLUID consists chemically of an aqueous solution of sodium chloride and bicarbonate containing mucus, albumin and *débris*. Its functions are to keep the surfaces of the conjunctiva and cornea moist, and to remove foreign bodies and organisms. The secretion of tears is increased (1) by irritants and foreign bodies coming in contact with the cornea, conjunctiva or lids; (2) by irritation of the nasal mucous membrane ; (3) by powerful illumination of the eyes ; (4) by the incidence on the eye of infrared (heat) or ultra-violet (actinic) rays ; (5) under the influence of emotion. When excessive tear formation occurs, the fluid either escapes over the front of the lids or is drained away through the lachrymal duct into the nasal sinus. Three theories have been advanced to explain the latter : (1) syphoning, owing to the mouths of the ducts being at a higher level than their exit into the nose ; (2) capillary, owing to the tendency of the liquid to flow into the ducts through surface tension ; (3) active removal by the act of blinking. It is not at the present time definitely known how this occurs. Some say that on closing the eye the internal palpebral ligament tends to be pulled on, and that this dilates and fills the lachrymal sac ; others that the sac fills automatically through having been previously emptied by the contraction of Horner's muscle. It is possible that both processes occur.

The eyes of some mammals and fish and nearly all birds are provided with a nictitating membrane, a semi-transparent shutter which can be brought over the surface of the cornea. It is supplied with plain muscle fibres innervated by the sympathetic. Stimulation of the sympathetic causes contraction of these fibres, resulting in retraction of the nictitating membrane, and so providing a ready means of estimating the degree of sympathetic activity in laboratory work.

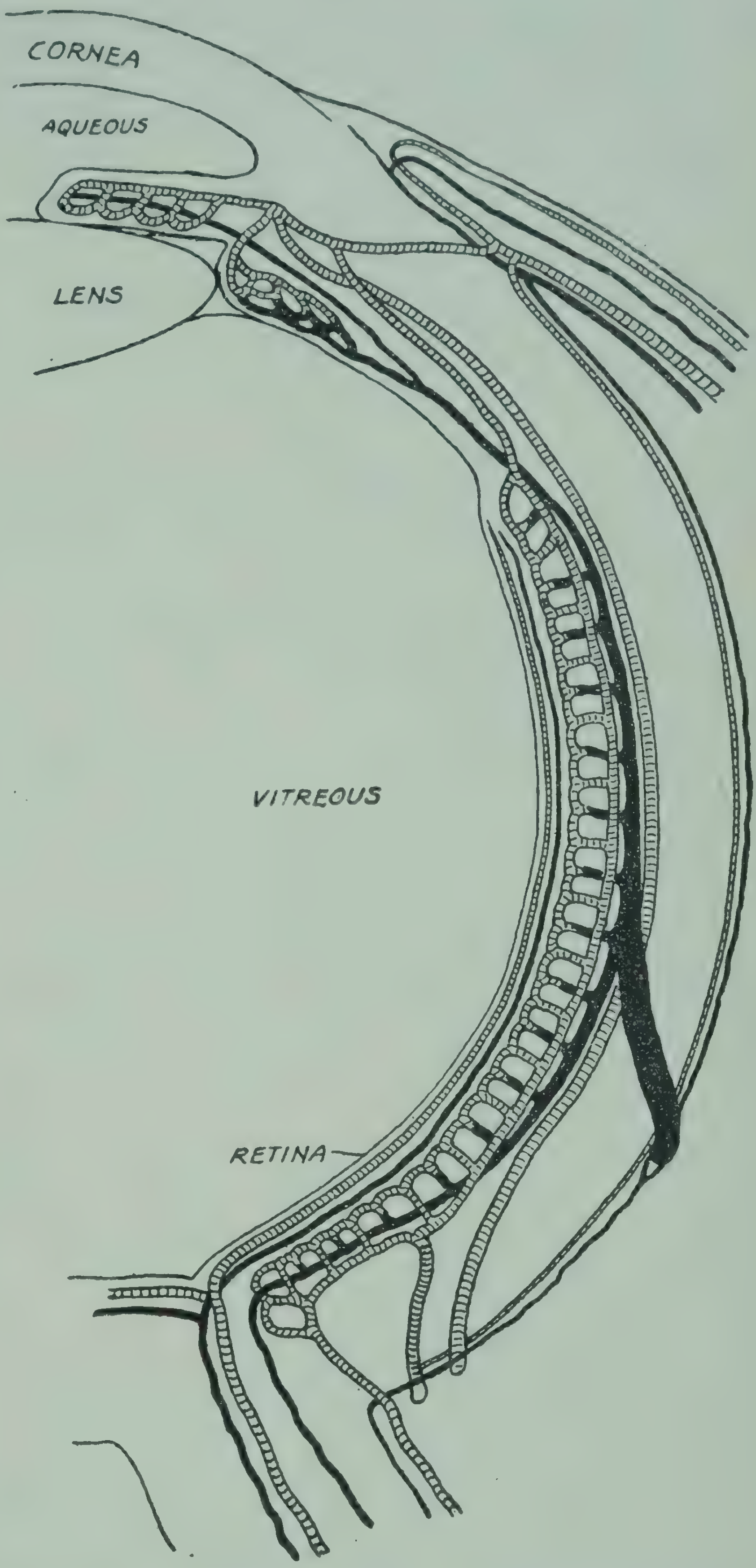


FIG. 248. Diagram to show the Blood Supply of the Eyeball. Arteries "lined," veins "black."

Nutrition of the Eye. The eyeball is richly supplied with blood vessels, which form numerous anastomoses. Among these may be mentioned the arteries of the optic nerve sheath, the long and short posterior ciliary arteries, the anterior ciliary arteries, which are branches from the muscular vessels, and the conjunctival arteries. These pierce the sclera to ramify freely in the choroid and in the ciliary bodies. The iris is supplied by two concentric vessels, the circulus major and the circulus minor. Between the two pass a number of radial fibres. The retina has a separate blood supply through the central artery of the optic nerve. Other structures, notably the transparent optical media of the eye, have no direct blood supply and therefore depend for their nutrition on diffusion from the aqueous humour. This fluid is formed principally by the ciliary bodies.

AQUEOUS HUMOUR. In chemical composition this fluid consists of water containing salts, traces of albumin and globulin, and a reducing sugar ; it is probably freely oxygenated. It has very nearly the same osmotic pressure as the blood (Duke-Elder). The fluid after secretion leaves the eye in one of three ways: (1) By travelling through the pupil into the anterior chamber of the eye and then through the spaces of Fontana at the edges of the iris (the so-called filtration angle) into the canal of Schlemm and thus into the ciliary veins. (2) Through the crypts in the anterior surface of the iris into the veins of that structure. (3) Between the suspensory ligaments of the lens to the anterior surface of the vitreous, then down the hyaloid canal to the papilla of the optic nerve, and thus out *viâ* the lymphatics of the nerve sheath or the retinal vessels. But whatever the fate of the liquid may be, it is certain that the amount secreted must be the same as that which leaves, because otherwise there would be a variation in the intraocular tension. Insufficient pressure will tend to disturb the correct relationship between the internal structures of the eye, and at the same time will prevent the proper action of the ciliary muscle in causing accommodation, because the suspensory ligaments of the lens will already be relaxed. Too great a tension on the other hand will interfere with the proper blood supply to the eye, and will prevent accommodation because the tension in the choroid will be too great for the ciliary muscles to overcome (see p. 389).

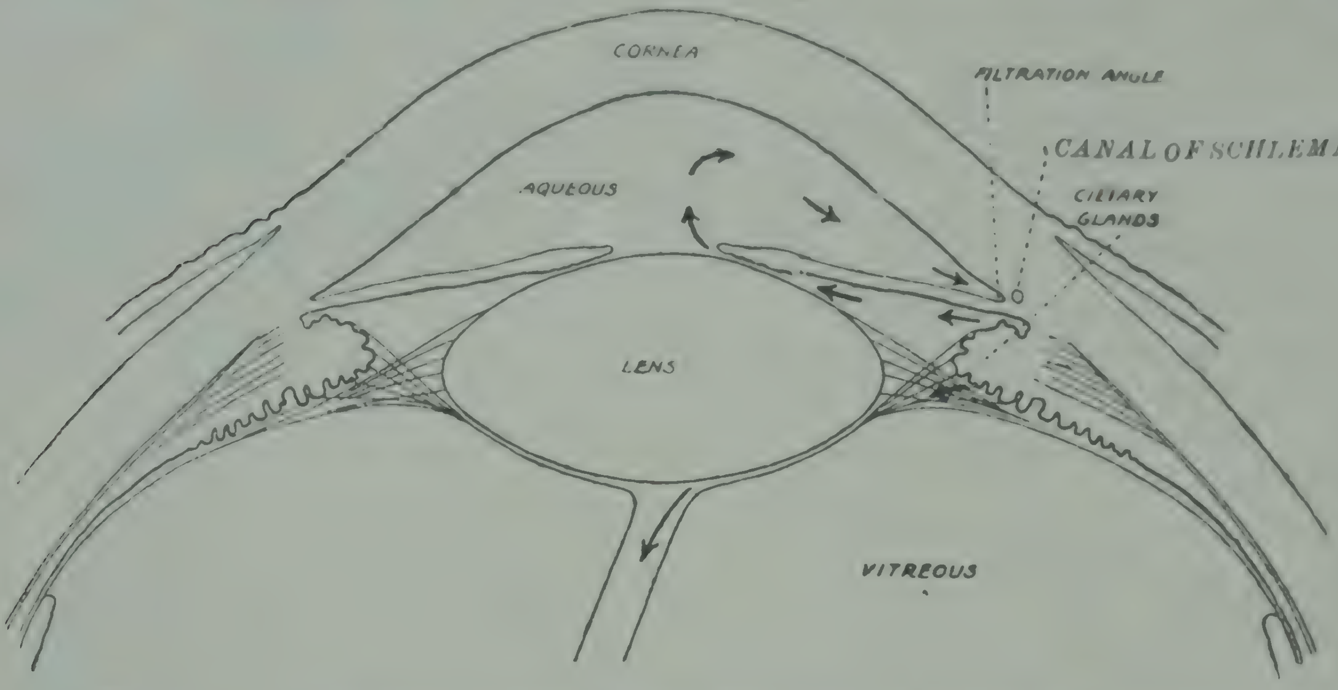


FIG. 249. Diagram showing Origin and Fate of Aqueous Humour.

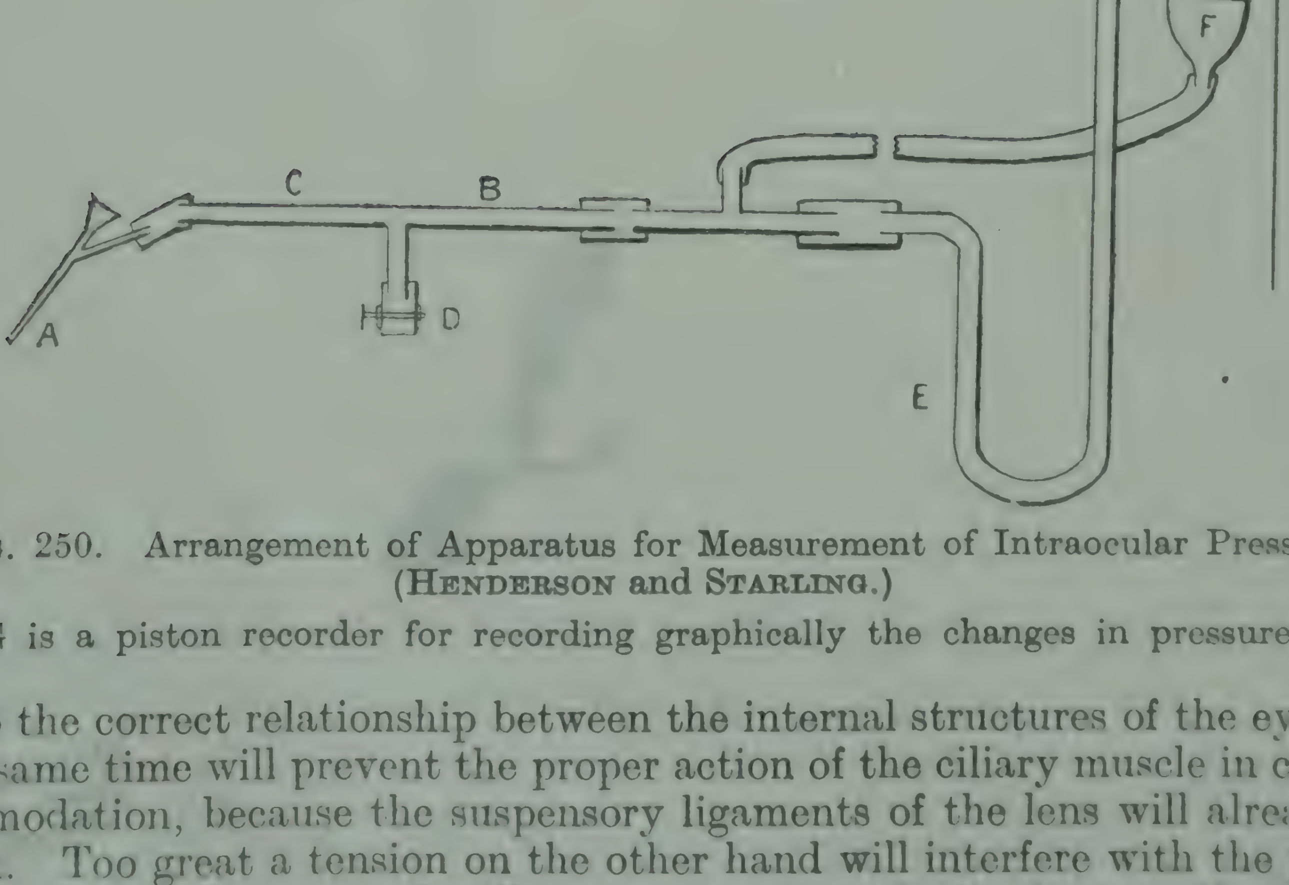


FIG. 250. Arrangement of Apparatus for Measurement of Intraocular Pressure. (HENDERSON and STARLING.)

G is a piston recorder for recording graphically the changes in pressure.

It is therefore important that there should be a proper control of the intraocular pressure. Experiments by Henderson and Starling, in which the intraocular pressure was determined by a null method (Fig. 250), showed that such a mechanism exists, because as the arterial pressure increased that in the eyeball increased much less.

Whereas the arterial pressure varied between 70 and 180 mm. (*i.e.* a difference of 110 mm.), the intraocular pressure was found to vary between 23 and 40 mm. (that is by only 17 mm.). The change in intraocular tension is therefore less than one-sixth of that taking place in the blood ; the control mechanism would therefore appear to have very considerable efficiency.

Glaucoma. The normal intraocular pressure in man is found to be between 25 and 30 mm. of mercury. The tension thus set up in the walls of the eyeball is principally borne by the sclera ; to some extent, however, assistance is rendered by the choroid owing to its elasticity, and by Tenon's capsule owing to the tonic contraction of its smooth muscle fibres (innervated by the sympathetic).

In abnormal conditions the efferent channels may become closed, either from pressure of the lens on the iris (as in hypermetropia) or from the presence of epithelial *débris* in the anterior chamber. The intraocular tension under these circumstances becomes very high, the disease being known as glaucoma. The principal symptoms of glaucoma are pain and impaired vision. The chief diagnostic signs are a stone-hard eyeball, sluggish, rather dilated pupils, and the retina when examined through the ophthalmoscope is found to show cupping of the optic disc, and vessels which are thin and show pulsation. In treating glaucoma, operative measures to lower the pressure should be taken immediately, because the high pressure interferes with the proper blood supply to the eye. Since all hypermetropes (persons with long sight) have a tendency to suffer from glaucoma, care should be taken against giving drugs such as atropine which causes dilatation of the pupil, since this increases the resistance to the escape of fluid at the filtration angle, and therefore predisposes to an attack of glaucoma.

For Reference

Duke-Elder (1934). Physico-Chemical Factors affecting Intra-ocular Pressure. *Physiol. Rev.*, **14**, 483.

Ibid. (1927). Recent Advances in Ophthalmology. (London : Churchill.)

4. THE OPTICAL SYSTEM OF THE EYE

The optical system of the eye consists of those structures which together focus an image of external objects on the retina. In the mammalian eye there are four concerned : the cornea, the aqueous humour, the crystalline lens and the vitreous humour. The histology of the cornea has already been considered. The aqueous humour is a structureless liquid. The vitreous humour which consists of a collagenous material is enclosed in the hyaloid membrane, and through the middle of it, during fœtal life, runs the hyaloid artery, which goes from the central artery of the retina to the posterior surface of the lens.

The Crystalline Lens is a biconvex transparent elastic body, enclosed in an elastic membrane called the capsule. To the periphery of the capsule are attached the suspensory ligaments of the lens, which are formed by the anterior radial fibres of the thickened portion of the hyaloid membrane (the zonula of Zinn) (Fig. 252). Between the suspensory ligaments are shallow pockets into which the ciliary processes fit closely. In this way the lens is held firmly in position, while at the same time, by the movement of the ciliary processes under the action of the ciliary muscle, the traction can be altered in the suspensory ligaments, thus effecting the change in curvature of the lens, which will be shown later to be necessary for accommodation.

Histologically the lens is composed of a number of radially arranged fibres, each of which is a modified epithelial cell. These fibres are arranged in concentric layers, the more peripheral being soft, nucleated and of low refractive index, while the central form a dense non-nucleated mass of high refractive index, the fibrous layers between having an intermediate structure and index.

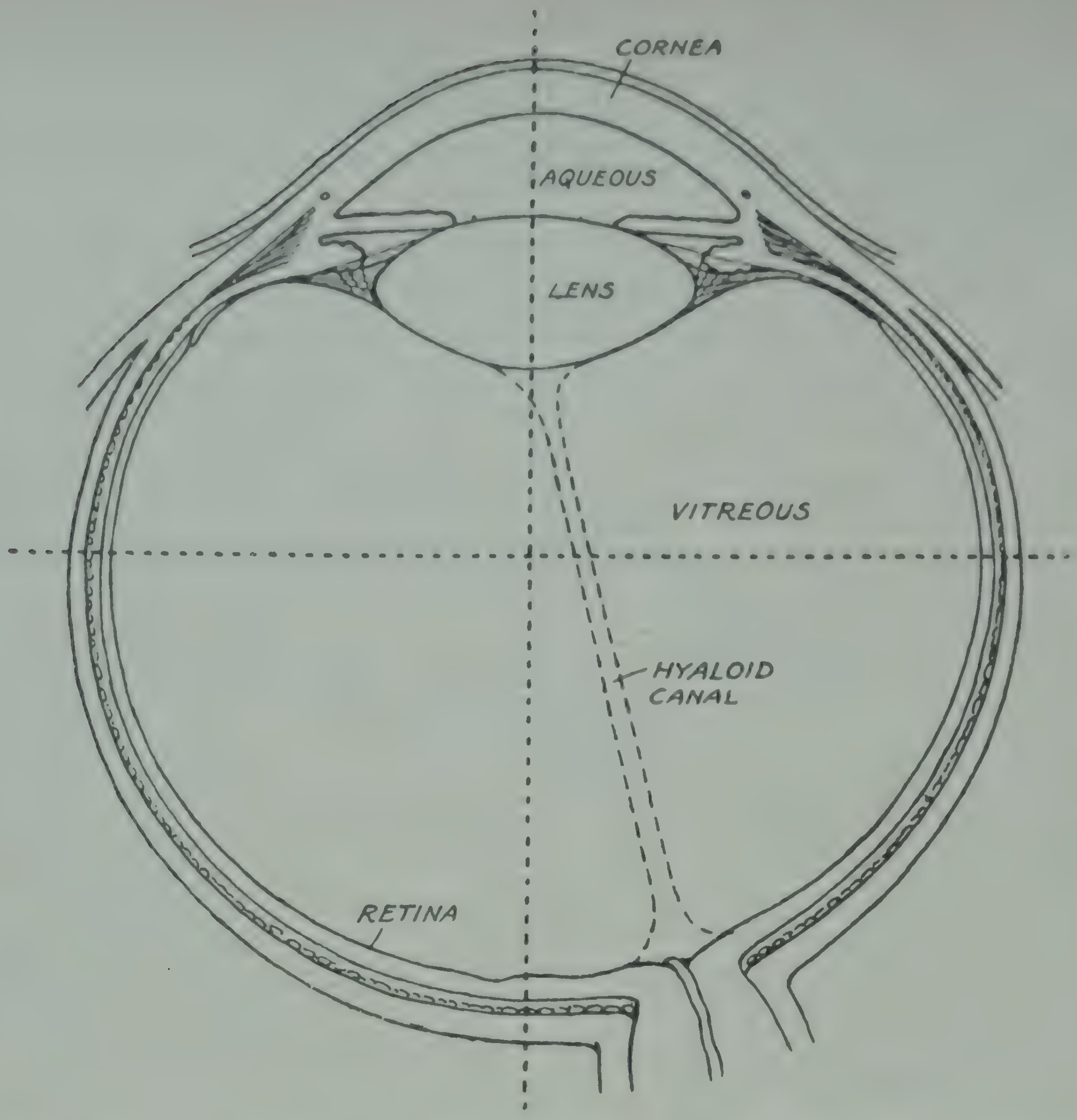

CORNEA
AQUEOUS
LENS
VITREOUS
HYALOID CANAL
RETINA

FIG. 251. Transverse Section through Equator of left Eye seen from above.

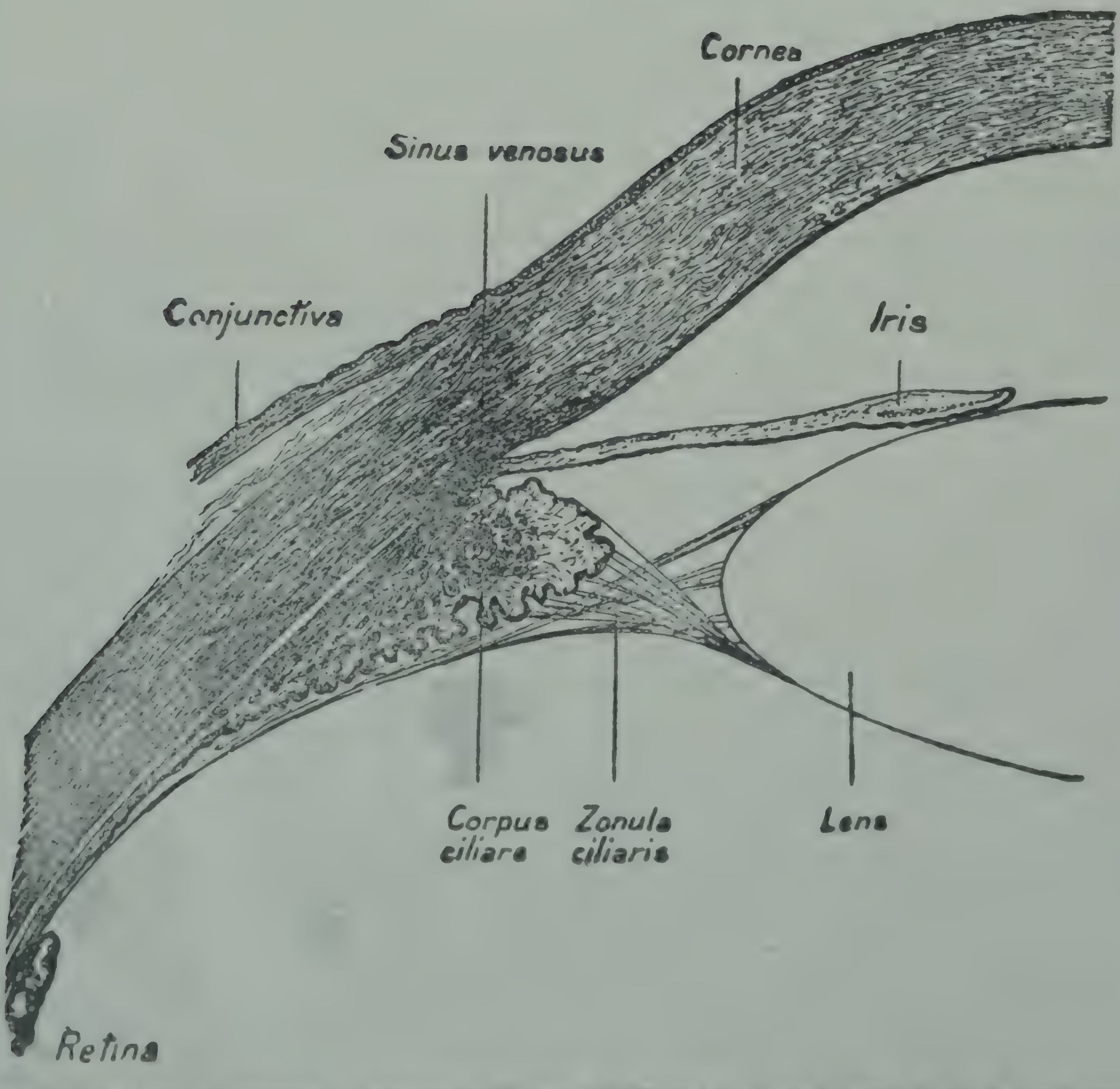

Cornea
Sinus venosus
Conjunctiva
Iris
Corpus ciliare
Zonula ciliaris
Lens
Retina

FIG. 252. Section through Anterior Part of Eyeball to show Mode of Suspension of Lens. (After MERKEL and KALLIUS.)

Metabolism of the Crystalline Lens. If the crystalline lens be removed from the eye of an animal and placed in a microrespirometer of suitable pattern, it can be shown that oxygen is taken up by the lens and that carbon dioxide is given out. If such a lens is deprived of oxygen, or is exposed to excess of carbon dioxide for some hours, it shows signs of degeneration. The lenses of young animals have a high cystine content, which is associated with an active glutathione oxidation system. As animals get older the cystine content falls, and this may be associated with the decreased rate of metabolism which the lens shows.

Refraction by the Crystalline Lens. Refraction occurs whenever light passes from one medium into another of different optical density. It is due to the fact that the waves, of which the beam of light is composed, travel more slowly in a dense medium than they do in one of less density. Some of the effects which this produces are shown in the diagram below.

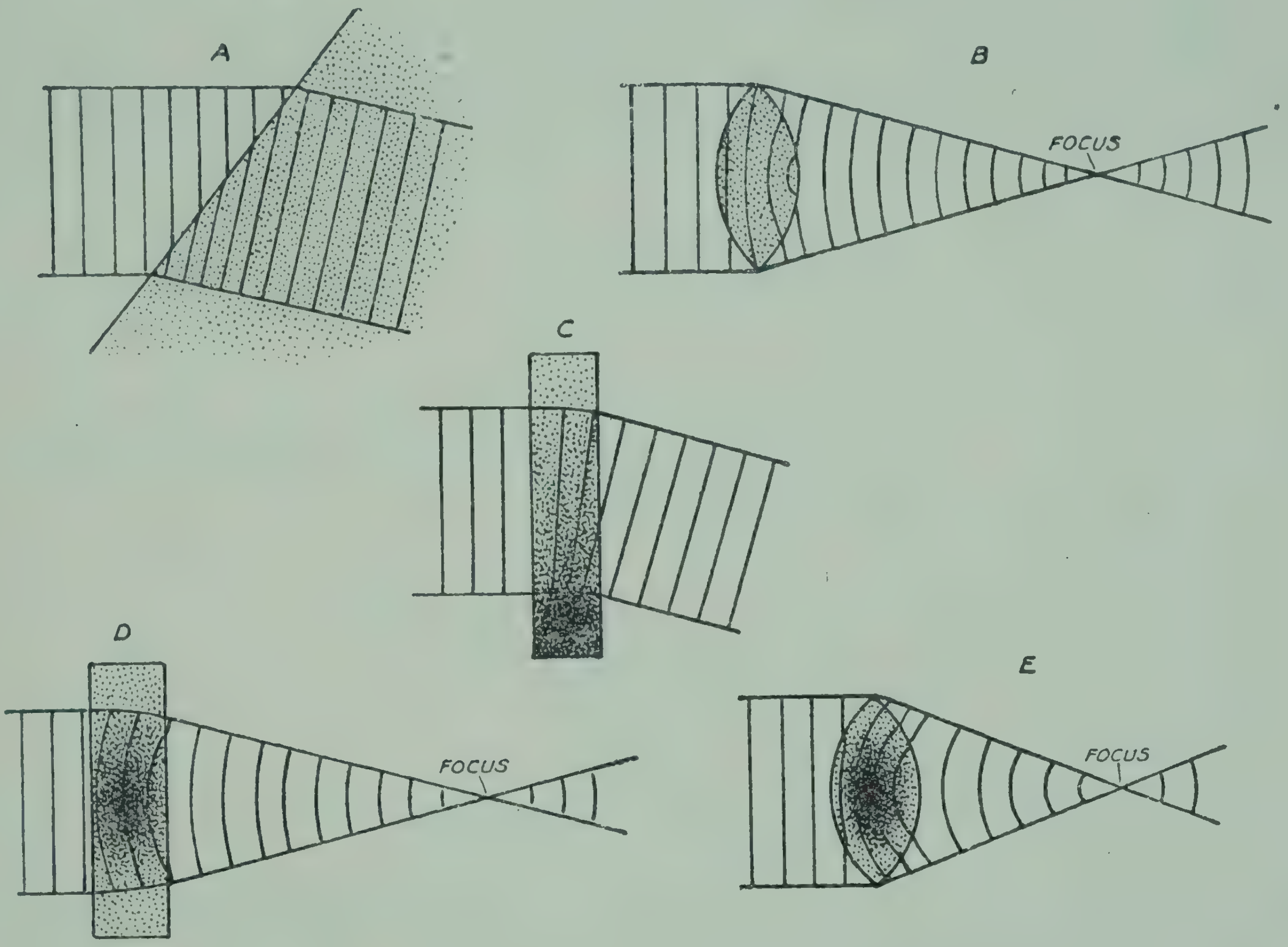


Fig. 253. Diagram showing Refraction of Light.

(A). By an inclined surface.
(B). By a lens.
(C). By a plate of greater density at its lower end than at its upper.
(D). By a plate of greater density at its centre than at its edges.
(E). By a lens of greater density at its centre than at its edges.

At A, plane waves are seen entering a dense medium at an angle. At B the medium is lens shaped. At C the medium has a plane surface but has a greater density below than above. At D the medium has a plane surface but a greater density at its centre than at its edges. At E the medium is lens shaped as at B, and also varies in density as at D. The very great refracting power of such a structure is well shown. This is the arrangement found in the lens of the eye, in which, therefore, the power of refraction is very much greater than that of an ordinary lens of the same curvature as the lens and of the same refractive index as the average density of its substance. The optical properties of the lens are therefore unique, and it is interesting to find the same (*i.e.* even immunologically identical) protein in the lens of different species of animals.

To show the effect which the increasing density of the lens produces, the refractive indices of its parts may be compared with its equivalent R.I. (that is, the refractive index of a glass lens of the same size, shape and focal length). The refractive index of the periphery of the lens is 1·37, and that of the central nucleus 1·41, the mean being about 1·39. But the equivalent density of the lens is found to be 1·42, that is, greater by 0·03 than the mean refractive index of its substance. The lens lies in contact with two transparent media, both of which have an approximate refractive index of 1·34. The power of the lens, if its composition were uniform, would therefore be proportional to the difference between its own mean R.I. (1·39) and that of its surroundings (1·34), that is to 0·05. Owing to its peculiar structure its equivalent R.I. is 1·42, and therefore its power is proportional to the difference between that and 1·34, that is to 0·08. Owing to its structure the lens has therefore increased in power in the ratio of 0·08 to 0·05. Now since the range of accommodation depends, other things being equal, on the power of the lens, we see that the peculiar structure of the lens has nearly doubled its range. The graduation in the densities of the different layers of the lens has a further advantage which will be described later, in that it reduces the spherical aberration of the eye as a whole and also reduces the amount of scattered light within the eyeball.

The Optical Constants of the Eye. In the case of the crystalline lens, two methods are available for the determination of the radii of curvature of the anterior and posterior surfaces, namely measurements on the excised lens, either in the air or preferably suspended in a fluid of known optical properties, or by estimating the apparent size of the images of an object which are formed by reflection at its surfaces. In order that the latter method shall succeed, a device must be employed for eliminating the effect of chance movements of the eye while under observation. This was first done by Thomas Young, who employed a method used in astronomy, namely that of doubling the image to be measured and then adjusting the lower edge of one image to be in coincidence with the upper edge of the other. If the eye moved during the determinations, both images moved together and therefore difficulties in adjustment were avoided. In the case of the cornea this method is alone available because only in the living state is the true curvature preserved (Fig. 254). In the case of the lens the determinations are complicated by the fact that the refraction of the cornea has to be allowed for. Further, the images that are seen are neither bright nor sharply defined; but in spite of this, considerable accuracy is attainable. The following are the approximate values given by these methods.

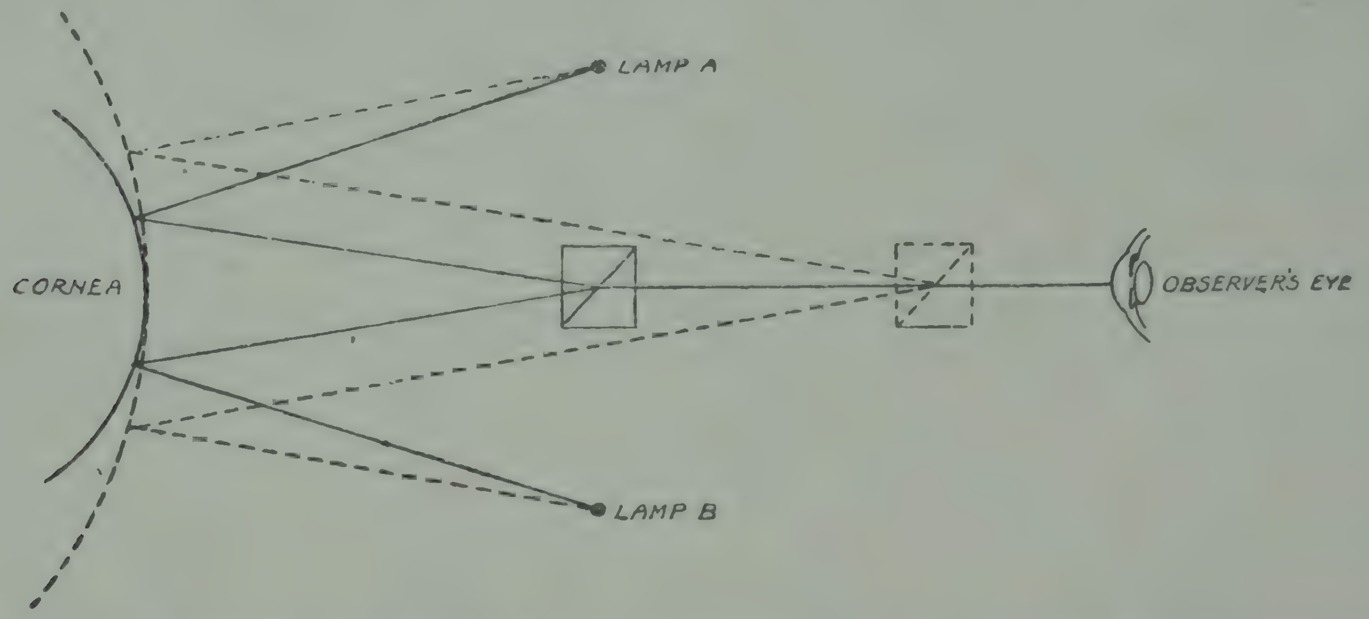


Fig. 254. Diagram to show a Method of Determining the Curvature of the Anterior Surface of the Cornea.

The images of the lamps A and B are caused to coincide by shifting the position of the double image prism. The greater the curvature of the cornea the closer must the double image prism be to the eye.

Radius of cornea	8 mm.
Radius of lens, anterior surface	10 ,,
Radius of lens, posterior surface.	6 ,,

The Refractive Indices (optical densities) of the eye media are determined on the excised eye by means of the Abbe refractometer. It is found that the cornea and aqueous are so nearly alike that for all practical purposes they may be regarded as one, particularly as the posterior corneal surface has nearly the same centre as the anterior. The refractive indices may therefore be given as follows :—

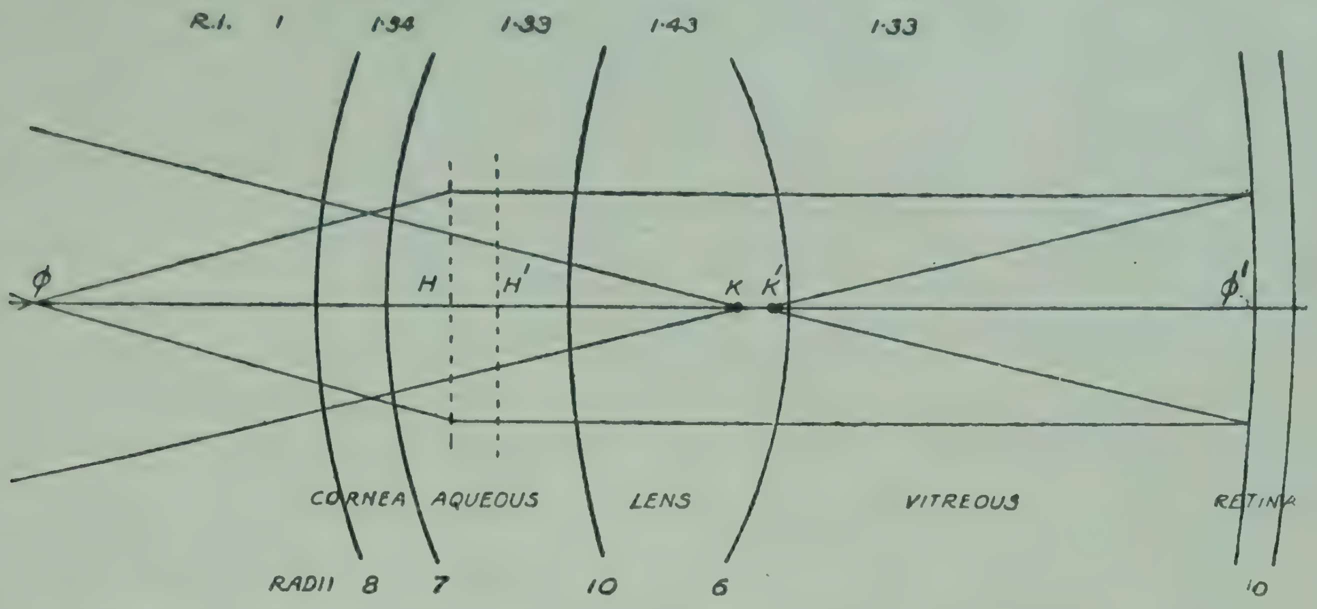


Fig. 255. The Optical System of the Eye shown diagrammatically.
H and H′ principal points. K and K′ nodal points. φ and φ′ anterior and posterior foci.

Refractive index of cornea and aqueous	1·34
Refractive index of lens (equivalent)	1·42
Refractive index of vitreous humour	1·33

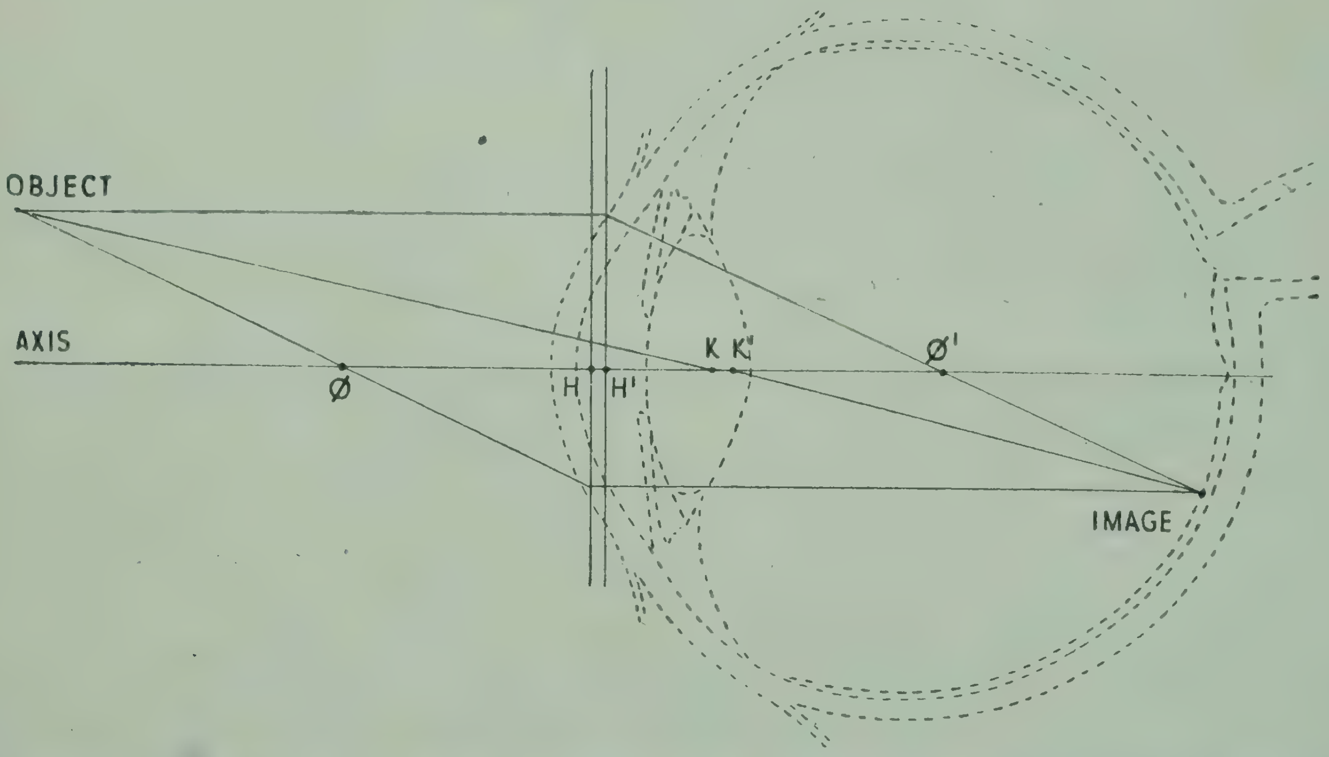


Fig. 256. Diagram showing approximately the positions of the principal planes and nodal points of the eye. The course of rays from a near object passing through the eye to form an image on the retina is shown.

The direction of the rays of light through the eye is shown diagrammatically in Fig. 256.

It is of some interest to know the relative part taken by the various refractive media of the eye in the formation of the image. By far the greater part is performed by the cornea. The following values in mm. and dioptres * may be given :—

	Mm.	Dioptres.
Focal length of the cornea	24	42
,, of the lens	44	23
,, of the whole eye	15·5	65

---

* In the dioptre nomenclature the figure represents the reciprocal of the focal length in metres. Thus 10 dioptres = $\frac{1}{10}$ metre = 100 mm.; and 24 mm. = $\frac{1000}{24}$ = 42 dioptres.

When opacities (a cataract) form in the lens this structure is removed by operation. It is then found that a lens of approximately 10 dioptres has to be worn by the patient in order that he may see distinctly. To this a cylinder has usually to be added to correct the astigmatism caused by the scar of the incision; and a 4-dioptre spherical lens has to be added when a glass for reading is being prescribed. If the glass lens could be placed inside the eyeball in the exact position occupied by the crystalline lens it would, as the above Table shows, require a power of 23 dioptres instead of 10 dioptres. When the eye is placed under water, the refraction of the cornea is necessarily abolished because water has approximately the same refractive index. Under these conditions the eye becomes too long-sighted for distinct vision. The eye of the fish has met this difficulty by the provision of a small, nearly spherical lens of very great density. In the fish, therefore, the lens takes the principal part on the refraction of the light in the formation of an image.

The Accommodation of the Eye. The above description has been made on the supposition that the rays entering the eye consist of parallel bundles, or in other words that the objects seen are at an infinite distance from the eye.

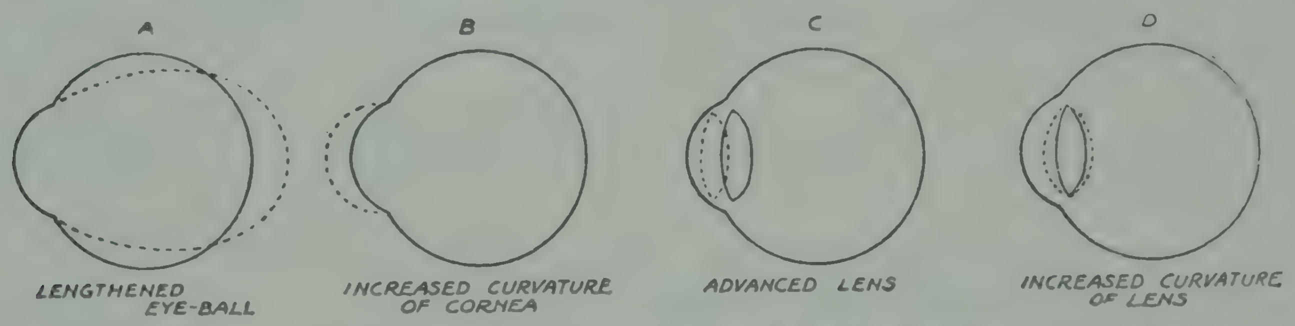


FIG. 257. Methods by which Accommodation of the Refraction of the Eye for Objects at Different Distances could be effected:

(A). By lengthening the eyeball. (C). By moving lens forward.
(B). By increasing curvature of cornea. (D). By increasing curvature of lens.

But during near vision, such is by no means the case. If there were no means of varying the focus of the eye, it would not be possible for divergent rays (those coming from near objects) to be brought to a focus on the retina. The mechanism for varying the focus of the eye is called accommodation.

THE THEORIES OF ACCOMMODATION. Of these, three are of historic interest only: (1) That during accommodation the cornea increases in curvature (similar to the bird's eye). This was disproved by Thomas Young, who placed his eye under water, replaced the corneal refraction by a convex lens, and then found the amplitude of accommodation unaffected. (2) That the eye elongates in near vision, thus causing the rays from near objects to focus on the retina (similar to the arrangement in the mollusc Pecten). This also Thomas Young disproved by placing two iron rings which could be clamped together, one in front of and one behind the eyeball. Having very prominent eyes he could do this if the eye being tested were rotated strongly inwards. The "phosphene," or subjective appearance, caused by the pressure of the posterior ring, which extended to the fovea, did not change in appearance during accommodation. He thus found no evidence for an elongation of the eye during accommodation. (3) That the lens during near vision advances towards the cornea (similar to the mechanism in the fish's eye). This view was disproved by Tscherning, who calculated that the lens would have to advance nearly 10 mm. in order to give the full amplitude found for the eye, whereas the anterior chamber of the eye is, as we have seen, approximately only 2·6 mm.

It follows that in the mammalian eye the fourth method is employed, viz. that accommodation is effected by increasing the curvature of the lens. As to the way in which this is brought about, there are three theories, those of Helmholtz, Tscherning and Thomson Henderson. The first theory (which is the one most generally accepted) supposes that the lens when removed from the eye is strongly convex and is accommodated for near vision. When in the eye, however, it is caused to become flatter through the traction of the zonula of Zinn (suspensory ligaments) on the edges of its capsule, and is therefore focussed for distance. But when the ciliary muscle contracts, it removes the tension on the zonula and therefore allows the lens to return by its elasticity to its more spherical form. Before describing the rival theories it would be well to examine the principal evidence on which Helmholtz' theory has been based. During near vision, measurement by means of the ophthalmometer shows that the anterior surface of the lens advances slightly and becomes at the centre of much greater curvature (10 mm. radius for distant vision, to 6 mm. for near). There can thus be no question that the change in the curvature of the lens is responsible for accommodation. The posterior surface is found to change but little; almost the whole range is therefore produced by the anterior surface. The changes found to occur in the lens of an adult twenty-five years old may therefore be summarised as follows :—

	Distance.	Near.
Radius of anterior surface, in. mm.	10	6
Radius of posterior surface, in mm.	6	5·5
Thickness of lens, in mm:	3·6	4
Focus of lens, in mm.	44	30
Focus of lens, in dioptres	23	33
Range of accommodation, in dioptres	10	

If the lens in near vision becomes more spherical owing to the relaxation of the zonula as Helmholtz supposed, we should expect a lens removed from the eye to be more spherical still, that is, in a state of strong accommodation. Tscherning stated that the changes in the curvature of the lens are much more complex than those given above. During accommodation not only is there, he says, an increase in the curvature near the centre of the lens, but at the same time a decrease in the curvature at the periphery. This view he supported by quoting the careful measurements which Young made by means of his optometer, and which have been confirmed by other observers. There is found to be a zone about 1·4 mm. from the centre of the lens where the curvature does not change appreciably. Inside this zone the curvature increases during accommodation, whereas outside the lens becomes flatter. Tscherning supposed that these changes of curvature are produced by an increase in the tension of the zonula during accommodation, in other words exactly the opposite action to that which Helmholtz supposed to occur. According to Thomson Henderson, the setting up of tension in the suspensory ligaments is due to the pull of the muscles of accommodation, that is, that these muscles are in tension during accommodation to distant vision. In the excised eyeball, therefore, the lens will be adjusted for near vision.

We may summarise the position with regard to these theories as follows :—

According to Helmholtz, the muscles adjust the lens for near vision. According to Tscherning and Thomson Henderson, they adjust the lens for distant vision. The state of the lens in an excised eye and when isolated from the eye is therefore as follows :—

	Helmholtz.	Tscherning.	Thompson Henderson.
Lens in excised eye accommodated for	Distant vision	Near vision	Near vision
Isolated lens accommodated for	Near vision	Distant vision	Near vision

Experiments have shown that the choroid moves forward during accommodation in both man and animals; further, Hess has shown that, when full accommodation has been performed, the lens is only loosely supported, so that gravity can act on it and cause it to sink slightly in relationship to the other eye structures These effects appear to be definitely in favour of Helmholtz' and Thomson Henderson's theories, and against that of Tscherning. Additional evidence on this point is provided by the experiments of Hartridge and Yamada (1922).

They obtained the following values in the cat's eye :—

Power of lens in excised eyeball	65 dioptres
Power of isolated lens	80 ,,

The lens in the excised eyeball is therefore in the condition for distant vision, whereas the isolated lens is in the condition for near vision, there being an increase in the power of the lens of 15 dioptres. These results are in favour of Helmholtz' theory and contrary to the theories of Tscherning and Thomson Henderson.

The Mechanism of Accommodation. The ciliary muscle consists of two separate sets of unstriated muscle fibres: the more superficial set of radial or longitudinal fibres, the deeper of bundles of circular fibres. The former take their origin from the sclero-corneal junction, and are attached to the anterior part of the choroid coat behind the ciliary processes. When these fibres contract, they draw the choroid forward and inward, the ciliary processes tending to occupy a smaller circle. The circular fibres lie in the substance of the base of the ciliary processes, so that, when they contract, they cause the apices of the processes to approach one another. In addition to these two sets of fibres, a third set has been described as meridional. These are, however, part of the radio-longitudinal set from which there does not appear to be much object in differentiating them. The ciliary muscles therefore have a common action, causing the ciliary processes to form a smaller circle.

The zonula of Zinn, or suspensory ligament, is formed of a large number of very fine fibres which run from the ciliary processes to the capsule of the lens. Further, those which arise posteriorly are attached anteriorly and *vice versâ*. When the ciliary muscle is in a state of rest, the tension in the choroid set up by the intraocular pressure causes the ciliary processes to be pulled in an outward and backward direction and therefore puts tension on the lens capsule through the zonula. The lens thus tends to be flattened and accommodated for distance. On contraction of the ciliary muscle the tension in the choroid is opposed and the ciliary processes approximated (Fig. 258). The zonula thus becomes slack, and the tension of the lens capsule decreases, allowing the lens to take up its more natural spherical shape, and thus to focus the rays from nearer objects.

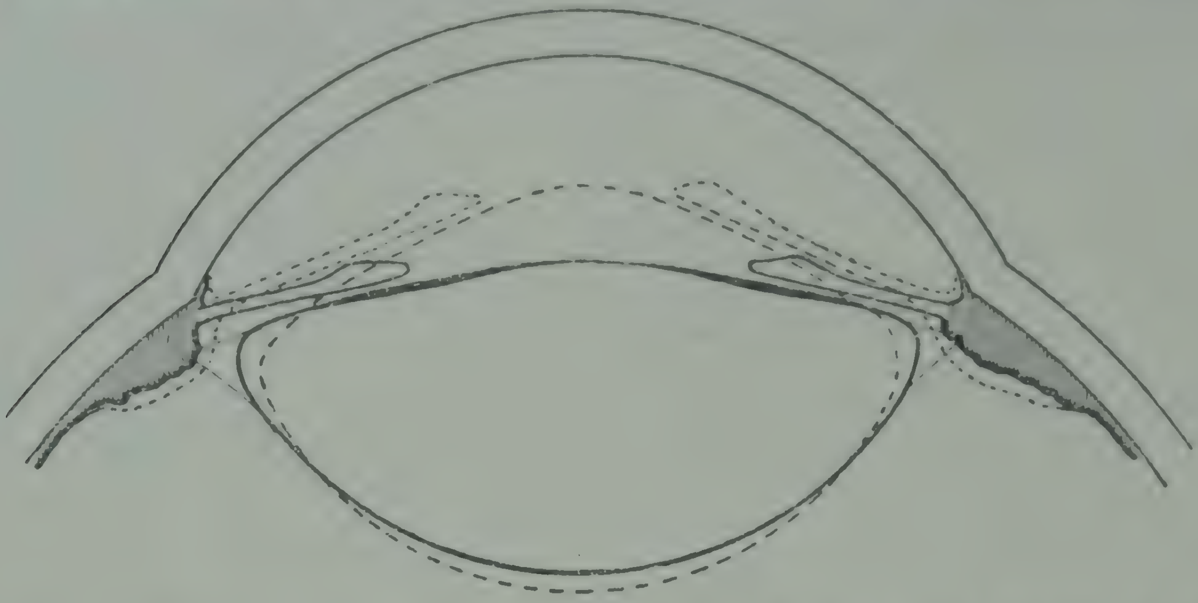


Fig. 258. Diagram to show the Changes in the Position and Shape of the Eye Structures during Accommodation. (After Fincham and Helmholtz.)

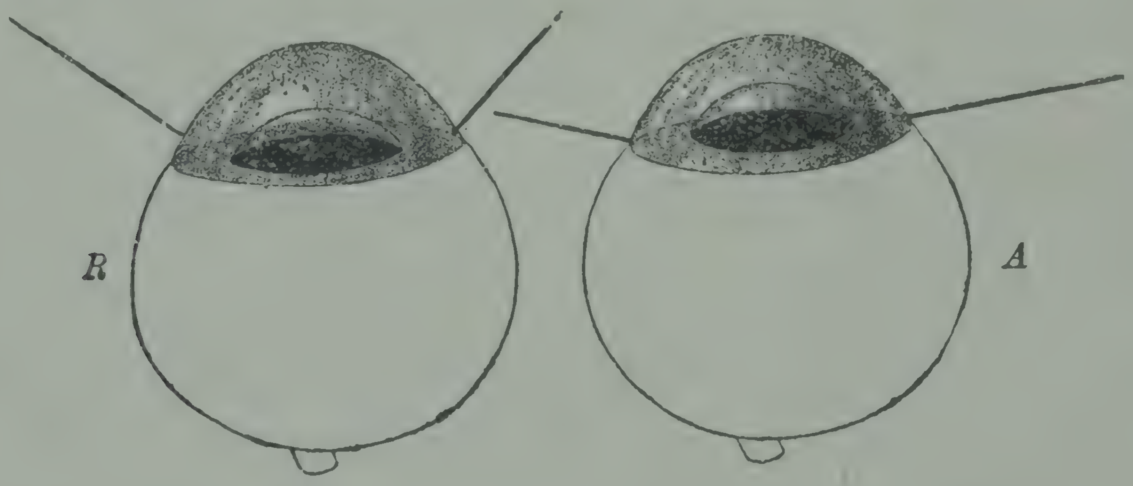


Fig. 259. Accommodation in the Cat's Eye. (Beer.)
R. Distance. *A*. For Near Vision.
Two needles have been passed through the edge of cornea into the ciliary bodies, to show forward movement of the latter during accommodation.

Fincham has pointed out the important part played by the variations in thickness of the capsule of the crystalline lens. This is very thin at the front surface near the centre, but very thick 2 or 3 mm. peripheral to the centre. The result of this is that the tension in the capsule causes the central part of the lens where the capsule is thin to become more highly curved, whereas it causes the more peripheral part of the lens where the capsule is thick to become flattened. These effects of the capsule are shown in Fig. 258 for both the unaccommodated and the accommodated lens.

Innervation of Mechanism of Accommodation. The ciliary muscle is innervated through the 3rd nerve, its nucleus being situated near the mid line under those of the pupils. Owing to the close association of the nuclei on the two sides, it is impossible to cause accommodation of the eyes separately. From this nucleus the fibres travel down with the rest of the nerve through the outer wall of the cavernous sinus, and when they reach the orbital cavity are given off to the ciliary (lenticular) ganglion, where they anastomose with nerve cells, the processes of which then proceed through the short ciliary nerves to the eyeball.

The position of the higher centre connected with the nuclei concerned in accommodation is not definitely known; it is believed to be in the occipital cortex. But since it is possible to carry out willed changes of focus as well as subconscious ones there must be connections with other parts of the brain as well.

THE AMPLITUDE OF ACCOMMODATION in the emmetropic (normal) eye is measured by ascertaining the nearest point from the eye at which perfect vision can be obtained. Since it is possible that the eyes, when examined separately, can focus nearer objects than they can when used together (owing to the limitation in the power of convergence), one of the eyes should be closed when making the determination. In the ametropic (abnormal) eye it is necessary to determine the far point as well as the near, since the former may not be at infinity as it is in the emmetropic eye. But other difficulties are encountered because, in the case of hypermetropic (long-sighted) eyes, an object placed at infinity still requires some accommodation in order to focus it. Lastly there comes the personal equation of the patient, since it is found that, even when apparently fully relaxed, the instillation of atropine usually causes some further relaxation of the accommodation; so in the same way the instillation of eserine is usually followed by a definite increase in the accommodation over that which can be voluntarily exerted. To obtain the maximum amplitude of the accommodation these drugs should therefore be used. Where a comparative and relatively approximate value is alone required they may be omitted. Of the many methods that may be employed, probably the simplest is by the use of the set of trial lenses, which vary in their curvature by small uniform amounts from strongly convex (plus) to strongly concave (minus). The test object consists of a pin placed vertically in a board at any fixed distance from the eye, a white surface being arranged behind it. A pair of spectacles to hold the trial lenses is placed before the eyes, and in them are inserted two metal plates with two vertical slits in them, so that each eye in turn may look at the test object through the slits in the plate opposite it. To determine the *near point*, minus (concave) lenses of gradually increasing power are placed before one of the eyes, the other being closed, until one is found that just causes the image of the pin to appear double. The power of the next weaker lens is therefore taken to be the correct one. To determine the *far point*, plus (convex) lenses of increasing power are tried in a similar manner, and that one taken which just does not cause appreciable doubling. The difference between the power of the two lenses found in this way gives in dioptres the value of the amplitude of the accommodation. Care should be specially taken to see that the value found for the concave lens has the minus sign placed in front of it. Thus if it was found that the far point was reached by the use of a convex lens of 2·25 dioptres, and that the near point required a concave of 7·5 D., then the amplitude is not 5·25 D. but 9·75 D. Careful measurements made in the above manner show that there is for different ages an average range of accommodation. In youth the amplitude is large, but it decreases uniformly to old age, and this decrease is called *presbyopia* (old-sight). The amplitude found as a rule at different ages is approximately constant and is given in the Table below.

Age in Years.	Accommodation in Dioptres.
10	13·8
15	12·6
20	11·5
25	10·2
30	8·9
35	7·3
40	5·8
45	3·7
50	2·0
55	1·3
60	1·1

This gradual reduction in the amplitude of the accommodation is caused by the hardening of the crystalline lens, which with the advance of years robs it of the elasticity of youth. As a rule the change goes on unheeded till the time comes when the near point of vision has receded so far from the eye that the book has to be held at a distance, in order to be focussed clearly. Then it is found that the print, particularly if it is fine, becomes unreadable, and convex glasses have to be worn. It is at that time that presbyopia may be said to begin, usually at an age of 45 to 50 in emmetropia. In hypermetropia, presbyopia shows itself earlier than normal, and in myopia (short-sight) later. The reason for these differences will be described when dealing with these conditions.

THE EFFECT OF DRUGS on the accommodation has already been alluded to. In ophthalmic practice three are used in order to paralyse the ciliary muscle, namely atropine (one-half per cent. solution), homatropine (2 per cent.) and scopolamine (one-fifth per cent.). One of these drugs is placed within the conjunctival sac, and from here it slowly travels, probably by diffusion *viâ* cornea and aqueous, to the ciliary and iris muscles, both of which are paralysed. This fact is of the greatest importance, for if this path did not exist the use of these valuable drugs would hardly be possible. Contraction of the ciliary muscles and of the sphincter pupillæ muscle is caused by the drugs pilocarpine and physostigmine (eserine). One of these is often used to counteract the effects of the atropine group.

5. THE REFRACTION OF THE EYE

In many respects the eye is an improvement on the ordinary photographic camera.

(1) The focussing of objects is performed automatically.

(2) The intensity of the illumination of the retina and its sensitiveness are adjusted automatically.

(3) The depth of focus is increased at the same time that the focussing of near objects takes place.

(4) The field of view is extremely large. In most cameras it does not exceed 90 degrees, whereas in the eye it is 208 degrees. Thus objects 14 degrees behind the observer are visible to him.

(5) In most camera lenses considerable loss of light occurs at the various surfaces of contact between glass and air. Thus in most photographic lenses there are four glass-air surfaces, and in many lenses the number is greater than that. In the eye, on the contrary, the space between the components is occupied by a medium not very different in refractive index from the refracting surfaces. There is thus only one surface where important reflection occurs, namely, at the anterior surface of the cornea. At all other surfaces the loss of light by reflection is extremely small so that the maximum light possible reaches the retina.

(6) The correction of the various aberrations from which a lens ordinarily suffers, are largely corrected in the case of the eye. In consequence its defining power is of a high order. The image of external objects which is thrown on the fovea centralis of the retina would hardly be improved if the lens system of the eye was removed and a microscopic objective of the same focal length substituted.

(7) The convergence of the eyes for the observation of near objects is automatically controlled.

(8) The retina consists of two receptor systems. One system is used for black and white in a dim light and the other is for colour used in a bright light. The fineness of grain in the latter is very good, exceeding 10,000 separate light receptors per square mm. The fineness is just adequate to make full use of the excellent definition of the lens system.

(9) The retina forms a cup-shaped surface which coincides everywhere with the position of the best definition of the lens system. Its curvature also causes the images of objects to have dimensions which are independent of their position on the retina. This factor is important for the provision of stereoscopic vision (distance perception).

Since the eye forms an image of external objects by means of its refracting media, it is found to have properties and to suffer from defects similar to those met with in the case of other optical systems. We may therefore treat the eye as if it were an optical instrument, and estimate its efficiency from that point of view. In the first place, therefore, we must consider what kind of an image it would form if it were a perfect lens system, suffering from no kind of aberration. Our experience of other lens systems, well-nigh perfect, has shown us that the image of a distant point source of light is not a mathematical point, as geometrical optics would have us believe, but is on the contrary a definite pattern of quite considerable size, the shape and dimensions of which can either be calculated from the conditions, or may be seen, measured and photographed by appropriate means. The formation of this pattern is due, not to any defect of the lens, but to the fact that light is a form of wave motion which exhibits a property called diffraction.

The diffraction pattern formed by any lens system varies directly as the focal length of the system and the wavelength of the light, and inversely as the diameter of the pupil through which the light passes. In the case of any lens system such as the eye, in which the pupil is circular, experiment and calculation are agreed that the image of a point source consists of a series of concentric rings of light, having a bright spot at their centre. The diameter of this spot in the case of the eye is found to be 0·01 mm. with a pupil of 2mm. diameter. No matter how perfect the eye be as an optical instrument, diffraction sets a limit in this way to the perfection of the image that can be formed. This should not, however, be thought of as a defect, but as a property, since it is caused by the nature of light itself. In the consideration of the principal optical errors of the normal eye, we have to decide in each case, not only to what extent the defect is present, but whether the defect produces any noticeable change in the diffraction pattern which may affect definition.

DEPTH OF FOCUS, like diffraction, is a property of a lens system and not an aberration. Its origin may be explained as follows :—Suppose objects 100 metres away to be forming sharp images on the retina, then objects at 200 metres will form images which come to a focus slightly in front of the retina, and objects at 50 metres images that are slightly behind. If however the focussed points are only a short distance in front of or behind the retina, the image of a distant point which fell on a single cone would still do so although its distance from the eye had been altered, because the cone has a certain diameter. Depth of focus in the case of the eye is the greatest distance through which a point can be moved and still produce an image which falls exactly on a cone without spreading at all on to neighbouring ones. For example, in the above case the distance moved was from 200 to 50 metres, that is, the depth of focus was 150 metres. Now it is found in the case of any lens system that depth varies with the aperture of the pupil. Thus in the case of the eye the following values are obtained :—

Pupil diameter.	Depth at infinity.	Depth at 25 cms.
1 mm.	From inf. to 8 metres	3·2 cm.
2 ,,	,, ,, 16 ,,	1·6 ,,
3 ,,	,, ,, 24 ,,	1·1 ,,
4 ,,	,, ,, 32 ,,	0·8 ,,

We see, therefore, that not only does depth decrease as the aperture of the pupil increases, but that it also decreases as the mean distance of the objects from the eye decreases. Thus with a pupil of 3 mm. the eye, if focussed sharply on objects 24 metres away, would also be in focus for objects at infinity and for objects at 32 metres. Depth of focus is therefore considerable at this distance. But if the eye is working at the distance of 25 cm., depth would be 1·1 cm. only. At a pupil diameter of 1 mm., depth would be increased threefold, and therefore the closure of the pupil, which accompanies accommodation and convergence for near objects, has the valuable property of increasing the depth of focus at the same time. The average reading distance depends on the size of the individual *i.e.*, on the length of the arm, and is about 35 cm. (14 in.) for ordinary adults.

CHROMATIC ABERRATION OF THE EYE. We have seen that white light consists of a number of rays of different wavelength, and that on refraction the short rays are more bent than the long. When, therefore, white light is

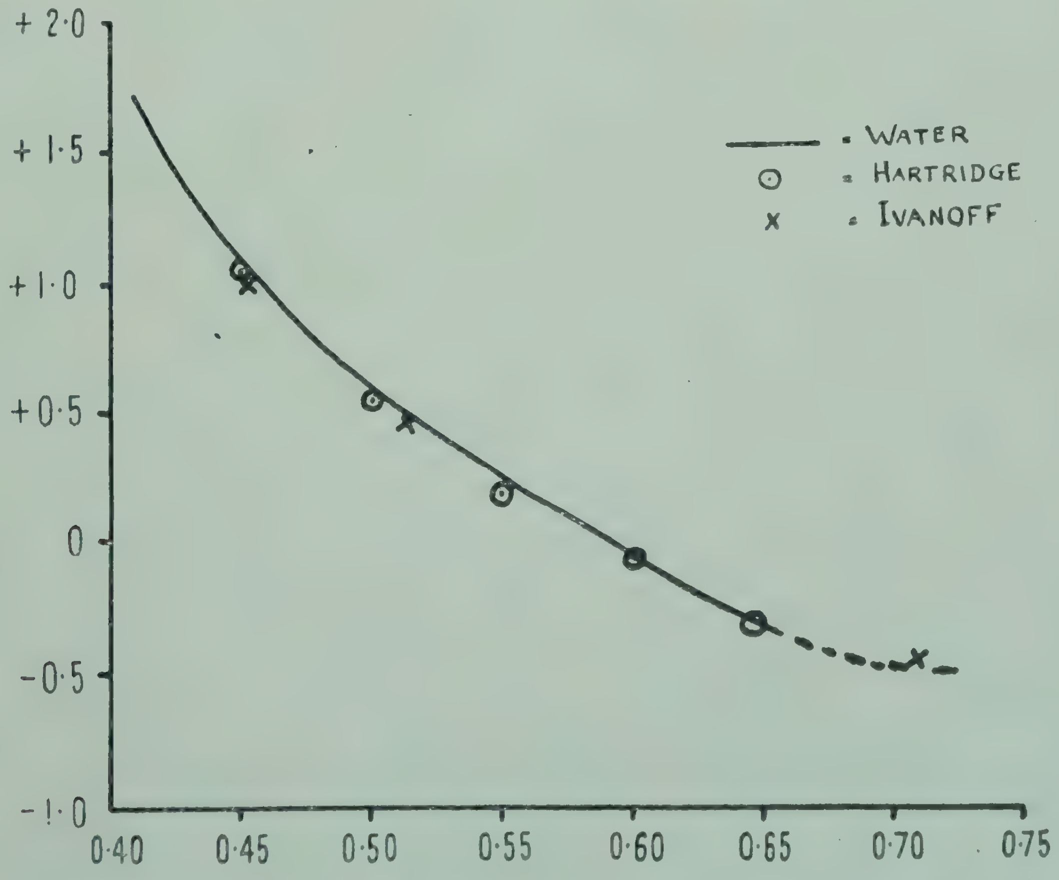


FIG. 260. Measurements of the Chromatic Aberration of the Eye. Refraction in dioptres vertical. Wavelength (μ) horizontal.

incident on a lens, the rays of short wavelength come to a focus in front of those of longer wavelength. This difference of focus for rays of different colour is called chromatic difference of focus (see Fig. 260). Experiment shows that, when such a series of foci are formed by the eye, the accommodation is so adjusted that the rays of greatest intensity (usually yellow rays) form the most sharply focussed image, and the colours of longer and shorter focus form blurred discs of light of relatively low intensity on top of this. Under these conditions it is found that quite well defined images are produced. Thus, with a pupil of 2 mm. diameter, approximately 70 per cent. of the light falls in an area of 0·005 mm. diameter.* Further it may be shown that a lens system such as the eye, which suffers from chromatic aberration, produces an image that is only just appreciably worse than one that is perfectly corrected, when the effects of diffraction are taken into account in both cases. But since the effects of chromatic aberration increase as the pupil enlarges, while those caused by diffraction decrease, it is clear that the larger the pupil the more does chromatic aberration tend to spoil definition. But as this is accompanied by a decrease in diffraction, the two changes taken together have the effect of leaving the actual definition practically unchanged. This important conclusion will be referred to again more fully in the last section. Besides effects on definition, chromatic aberration causes small bright points of light on a dark ground to form images which are largely composed of yellow rays, and on the other hand small black objects on a bright ground to be violet in colour. The reason for these colours being unnoticed in ordinary circumstances is due to the presence of the complementary colour which forms a fringe round the central point.

What happens is this: with a reduction of the visual angle, which is subtended by a yellow test object at the observer's eye, the yellow colour is replaced by white or by pale grey. Under similar circumstances blue or violet are replaced by dark grey or black. But experiment brings out the interesting point that both yellow and blue test objects tend to lose their colour more readily when they are in contact than when they are separated; thus a grating test object consisting of alternately arranged yellow and blue bars is easily mistaken by the observer for one consisting of black and white bars. Similar losses of colour are found to occur in the case of the pairs of fringes which are produced by the chromatic aberration of the eye. This subject will be found referred to in greater detail on p. 444 in the section describing "The correction of the chromatic aberration of the eye."

SPHERICAL ABERRATION OF THE EYE. The employment of spherical surfaces as boundaries for optical media leads to a difference in the position of the foci of rays that have passed through the centre of the lens and those that have passed through the more peripheral parts. The latter usually form a focus nearer to the lens. Since the eye is bounded by nearly spherical curves, it might be assumed that this aberration must be present in this organ. It should be remembered, however, that the crystalline lens has a structure quite different from that found in the lens systems of optical instruments. For the presence of a graduation of optical density, culminating in a nucleus of relatively great curvature, causes rays passing through the centre of the eye to be refracted to a greater extent than more peripheral rays, or, in other words, exactly the opposite effect to that produced by spherical aberration. Measurements on the eyes of different individuals, therefore, show the presence both of small amounts of under-correction (when the correcting effect of the lens nucleus has not been enough) and also actually of over-correction (when the lens nucleus has had too big an effect). In quite a number of cases the amount of spherical aberration is negligible even with pupils of 4 mm. diameter. With larger pupils there is probably a certain amount of under-correction, but this again is less than would be found in the case of spherical surfaces, because the more peripheral parts of the cornea are flattened and therefore refract less (as

---

* HARTRIDGE. *J. Physiol.*, 1918, **52**, p. 175; and 1922, **57**, p. 52

Gullstrand has shown). We may say, therefore, that in everyday life the effects of spherical aberration are altogether negligible, compared with those of diffraction and chromatic aberration.

PERIPHERAL ABERRATIONS OF THE EYE. So far the definition of an image lying on the principal axis of the lens has alone been considered. When this is not the case, other conditions are encountered which are more unfavourable in their effects. In the first place, the rays that form images on the peripheral parts of the retina make considerable angles with the surfaces of the eye media. This will cause chromatic difference of magnification, since blue rays will be more bent and will therefore form smaller images than red rays. It will also introduce " comma," that is, the effect due to non-compliance with the sine condition. It is seen at once, therefore, that the image formation by the periphery of the eye is altogether more imperfect than it is at the centre. The presence of the nucleus of the lens still further impairs the marginal definition. In fact we may say that in the eye, as in the microscope objective, the marginal images have been sacrificed in order thereby to improve the central ones. That this has been a very valuable result will be shown later. It will be shown in the next section that the most sensitive region of the retina is not exactly in correspondence with the optical axis of the lens system of the eye, being displaced approximately 0·5 mm. to the temporal side.

STRUCTURE OF FOVEAL IMAGE may be determined approximately by considering in turn the effects of the different aberrations on the light rays which enter the eye. For this purpose the only errors of importance are chromatic differences of focus and magnification. In addition, however, we must take into account the very important effects of diffraction. The final results of such a calculation show that the images of rays of different wavelength overlap one another. At the centre of the image is seen the sharp yellow focus of the highest intensity. Eccentric to it, and overlapping one another, are seen the diffuse red and green foci, which are of much less intensity. Where these overlap they produce a compound yellow according to the rules of colour mixture. Further outwards is the still more diffuse image of the blue rays, which is of almost negligible intensity. It is seen therefore that the centre of the image is entirely occupied by the sharp and intense focus of the yellow rays. Not only are these rays the brightest in the spectrum, but they are also those nearest to white light in their physiological properties. It is because of this structure of the image that the acuity of vision is so great at the fovea.

PERIPHERAL IMAGES have, as stated above, been to a considerable extent sacrificed, so far as their definition is concerned, in order to obtain the best possible conditions at the fovea. We find, therefore, at the periphery images that in no way compare with those formed near the optical axis of the eye. Even here, however, there is evidence that the eye has been designed to give the best results obtainable.

Scattered Light. In describing the histology of the retina it will be shown how generously the layer of cells lying immediately outside the sensitive layer of rods and cones is supplied with pigment; the object of this is clearly to absorb scattered light. In spite of this, however, we find considerable amounts of light being reflected back again by the retina; in fact it is this light that enables us to see the retinal nerves and vessels through the ophthalmoscope (see p. 411).

Abnormal Refraction of the Eye. It would almost be anticipated that such a complicated organ as the eye would be found to show individual abnormalities. A further consideration would probably suggest to us that, considering the smallness of the change that is necessary in any one of the optical media in order completely to destroy definition, it is nothing short of astonishing that abnormality of refraction is relatively so uncommon. In the newly-born the eye is almost always long-sighted (hypermetropic); this is due to the eyeball being too small for the optical system which it contains; the image formed by the latter is therefore focussed behind the retina. As age advances the eyeball grows until the point is reached at which the eye is emmetropic (normal). If, however, the child is allowed to use its eyes too much for near work, the eyeball goes on increasing in size until it has overshot the mark, and has thus caused the eye to become short-sighted (myopic). There would appear to be some kind of automatic control, which causes the eye to grow till it is in adjustment with the conditions most frequently encountered. This hypothesis is confirmed by the fact that if a child which is beginning to develop short-sight is prevented from using near vision for a year or two, the development of short-sight stops. The importance of the early detection of the onset of short-sight therefore cannot be too strongly urged.

The Methods of Diagnosis. The detection of errors of refraction in the eye may be effected in various ways, each of which is said to possess advantage. Some of these have come into such general use that they may be briefly considered as an introduction to the description of the more important types of error which they are used to investigate.

The Determination of the Visual Acuity. It has been found by experiment that persons with normal sight can distinguish between objects when the angle separating them is not much less than one minute. Test type has therefore been prepared in which the letters are composed of lines which subtend this angle at the eye, when the type is placed at a standard distance of six metres. Persons who are able to read the type at this distance are said to have normal vision. Above these standard letters are placed a series of larger letters, which at two, three or four times the standard distance would subtend the standard angle. A person with reduced acuity might be able to read at six metres the type that should be read at sixty. He therefore has visual acuity one-tenth the normal. Such a person might have long-sight, short-sight or astigmatism: to determine which is present, there is placed before his eyes a lens-frame, into which can be inserted any two of a large selection of glasses of different power, which are known as trial lenses. These are tried in turn in an orderly manner, until some are found which allow the man to read the standard type at the standard distance. His visual acuity is now at the normal and the strength and shape of the glasses in front of his eyes is carefully noted, so that others of the same power may be fitted to spectacles for him to wear. If the glasses are found to be convex (plus), then he was suffering from long-sight (hypermetropia), and if concave (minus) from short-sight (myopia). Very great care must be taken in this latter case, however, not to prescribe a stronger lens than is necessary. But if, on the other hand, cylindrical lenses had to be used, then he had astigmatism, either alone or in conjunction with long- or short-sight.

The Method of Retinoscopy. Such a method as that just described could only succeed if the person tested were an intelligent adult, because we depend entirely on his giving the correct answer, when we ask if the substitution of a different lens for the one we have already placed before him makes vision better or worse. With a child such a method could never succeed. Another method is therefore practised, which has the great advantage of being independent of the patient; in fact for the purpose of the test he might be blind. This method consists in throwing, into each of his eyes in turn, a beam of light reflected from a plane mirror, in the centre of which is a hole through which the observer looks. When the beam of light is directed into the patients' eye, the observer sees through the pupil a pink reflected beam of light. As the mirror is gently tilted, so as to throw the beam slightly upwards and slightly downwards, so the pink beam appears to move up and down behind the patient's pupil. If it moves down as the mirror is tilted down, the movement is said to be WITH the mirror, and the patient is hypermetropic, requiring plus spectacles. If on the other hand the beam moves "AGAINST the mirror," minus spectacles are required, since the patient is short-sighted. By placing glasses of different power in front of his eye, until one is found which causes the pink beam to move neither with nor against the mirror, the actual power for the spectacles required by the patient is ascertained. It should be carefully noted however that, since the observer is standing at about a metre distance from the eyes of his patient, plus one D must be subtracted from the power of any glasses that are found to be necessary. Thus if the patient was found to be myopic and minus 7 D spherical lenses were required to neutralise the movement of the beam, then the power that should be ordered is minus 8 D spherical. This test is found to work admirably in practice. It is usual to paralyse the pupil reflex and the accommodation of the patient previous to the test by the use of atropine, but some say that this is unnecessary.

Other Methods. Of other methods of testing vision little need be said: some require the use of special instruments such as the optometer and the refractometer. Another again depends on the determination of both the far and near points. This is of distinct value because it at once gives the amplitude of the accommodation, which is an important determination. Others are based on the use of the ophthalmoscope. But none of these methods is so simple or accurate as the method of retinoscopy described above.

Stenopæic Aperture. Often in practice the question arises as to whether low visual acuity is due to defect in the optical media of the eye, or to disease of the retina. This question can be readily answered by placing in front of each eye in turn a metal disc in which has been drilled a hole 0·3 mm. in diameter. If this improves acuity the defect is not in the retina; if it does not, it is. This test should be done in a good light because of the small amount of light passed by the hole. A hole used in this way is called a stenopæic aperture.

Hypermetropia or Long-sight. There are two principal varieties of long-sight: firstly, that in which the eyeball is too small and too short for the normal optical system; secondly, that in which the eyeball is normal but the refracting power of the lens below the normal. The first variety is found in childhood, because the optical system reaches its adult size much earlier than does the eyeball. In the majority of children the eyeball

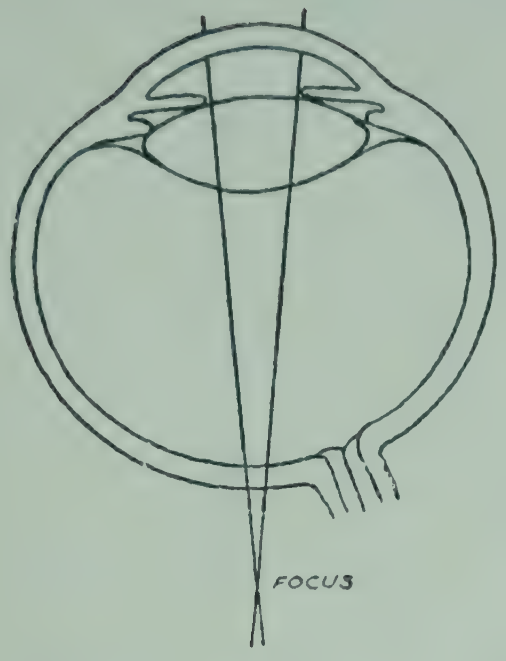


FIG. 261. Hypermetropic Eye.

The eyeball is too short, and therefore rays from a distant object come to a focus beyond the retina.

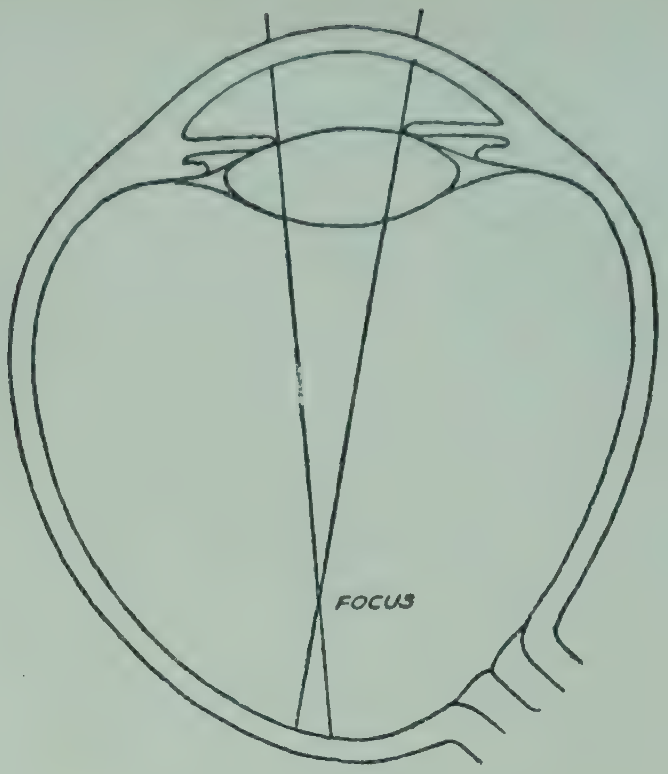


FIG. 262. Myopic Eye.

The eyeball is too long, and therefore rays from a distant object come to a focus in front of the retina.

continues to grow until it is the correct size, and therefore long-sight disappears. In a certain number of cases this does not happen, and therefore long-sight remains through life. The second variety is found in old age, and appears to be due to the absorption of water by the lens. The result in both cases is that the rays of light from distant objects are brought to a focus behind the retina, and therefore in order to focus them the accommodation has to be used (see Fig. 261). It follows from this that there is less accommodation remaining for the focussing of near objects, and therefore an inability to see distinctly at relatively short distances from the eye. Thus the use of the term long-sight.

Hypermetropia in adults is, therefore, more an abnormality than a disease; it causes a disposition, however, to three more serious conditions, namely glaucoma, internal strabismus and eye-strain. Glaucoma has already been described (see p. 383); it is due to an abnormal rise in the intraocular pressure, which occurs owing to the free escape of the aqueous humour at the filtration angle being checked. Now in hypermetropia we have seen that the eyeball is too small for its optical apparatus, and therefore the lens occupies too much of the space in the small anterior and posterior chambers. This causes the ciliary bodies and roots of the iris to be squeezed and greatly reduces the space at the filtration angle. An attack of glaucoma is therefore more liable to occur in the hypermetrope than in a person with normal refraction.

Internal strabismus is caused in hypermetropia by the accommodative effort that is made in order to focus an image on the retina because, as we have seen above (p. 374), convergence and accommodation are associated actions. When, therefore, the long-sight has been corrected by means of spectacles, and the accommodation is no longer called into play for seeing at a distance, the associated convergence no longer occurs and the strabismus disappears.

Eye-strain is caused in hypermetropia by the continual call for accommodation. Further, this must occur without convergence, for otherwise diplopia (seeing double) and strabismus develop as just described. A special strain is therefore placed not only on the ciliary muscles, but also on the external eye muscles. This state of affairs very rapidly causes fatigue, and headaches are therefore common.

The treatment of long-sight consists in prescribing suitable convex spectacles. It should be noted that the amount of long-sight actually present is shown only when

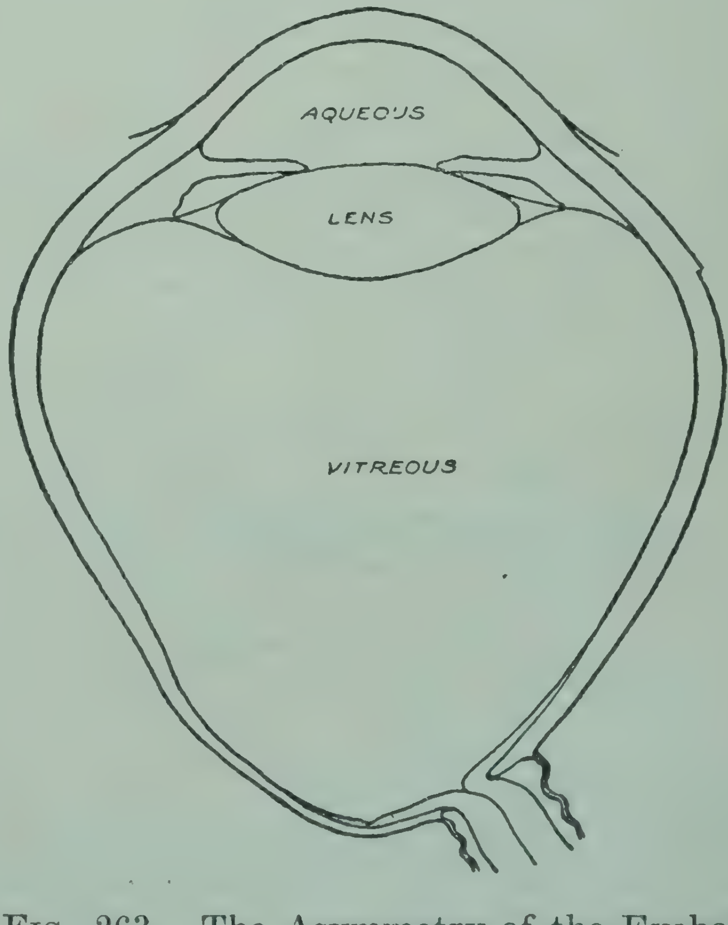


FIG. 263. The Asymmetry of the Eyeball and Kinking of the Optic Nerve caused by High Myopia.

the accommodation has been paralysed by atropine, because the patient has grown so accustomed to use his accommodation in ordinary vision that he is unable voluntarily to relax it. There is a certain amount of spasm of the accommodation. Because of this, the glasses prescribed should be less strong at first than the full correction shown to be necessary. These may be substituted by more powerful ones later.

Myopia or Short-sight. In this condition parallel rays, that is those coming from distant objects, come to a focus so far in front of the retina that the image appears blurred (see Fig. 262). Myopia may be caused in two ways, which are similar but opposite in action to those that cause hypermetropia; the first type is caused by the eyeball being too long, and the second by the refraction of the lens being too high. The former, which is the more common, usually develops in youth, particularly at the school age when the growth of the body makes special demands on the system, and at the same time feeding is usually bad. The constant use of the eyes for near work causes them considerable strain, which they are unable to withstand owing to their being unable to compete for nourishment with the rest of the body. The choroid and sclera therefore become thin, are no longer able to stand the tension set up by the intraocular pressure, and therefore expand, causing the eyeball to become larger than normal, and taking the retina beyond the focus of the optical system (Fig. 262). The treatment of myopia is therefore not only the wearing of spectacles, but the absolute prohibition of near work or close study, the administration of extra-nourishing food and an open-air life for a year or more. If these steps are taken at once, the myopia may get no worse, and may in fact get better. But if neglected the condition will almost certainly get worse. As myopia is a disease particularly liable to occur at the school age, schoolmasters and others associating with children should be on the look out for conditions likely to cause it, such as bad light, bad food and poor ventilation, and for its presence in any of the children. Glasses should always be prescribed, and care taken that the child wears them constantly, because it is found that beside assisting good definition and relieving eye-strain, they actually tend to check the further development of the trouble.

Certain complications sometimes attend myopia; these are divergent strabismus, eye-strain, and spasm of the accommodation.

The divergent strabismus has a similar origin to the convergent strabismus met with in hypermetropia, namely association of deviation of the eye axes with the adjustment of the accommodation. Now since, in the normal individual, the relaxation of the accommodation of the eye is associated with parallel axes of the eyes (in order to look at distant objects), in myopia the disuse of the accommodation for near vision causes the eye axes to remain straight and therefore produces the effects of an external strabismus. The use of glasses introduces again the necessity of accommodation, exactly as if the eye were normal, and therefore abolishes the strabismus. In the majority of cases an actual strabismus does not develop, but there is nevertheless a strong tendency to diplopia, especially when the eyes are tired. The eye-strain which frequently accompanies myopia probably has its origin in the effort to converge the eye axes without at the same time calling the accommodation into play.

Spasm of accommodation frequently accompanies myopia, and has the effect of making the myopia seem greater than it actually is. The true state of affairs is at once found when atropine is used, because the accommodation is thus abolished. Sometimes in children spasm of accommodation occurs without any actual abnormality of refraction. Such cases should be treated with the same care as those that are already developing myopia.

Astigmatism. The condition of the eye called astigmatism is one in which parallel rays are not brought to a focus in a single plane, but in a number of different planes. There are two different varieties of astigmatism. In the first, or irregular, variety the separate parts of one meridian of the eye form different foci. This is found to occur during the development of cataract in the crystalline lens, and also after ulceration of the cornea. The effects of this form of astigmatism on vision vary with the severity of the condition; in moderate cases a frequent phenomenon is the formation of a double image in the affected eye. Glasses as a rule do not give benefit. In severe types the use of a stenopæic aperture may improve definition.

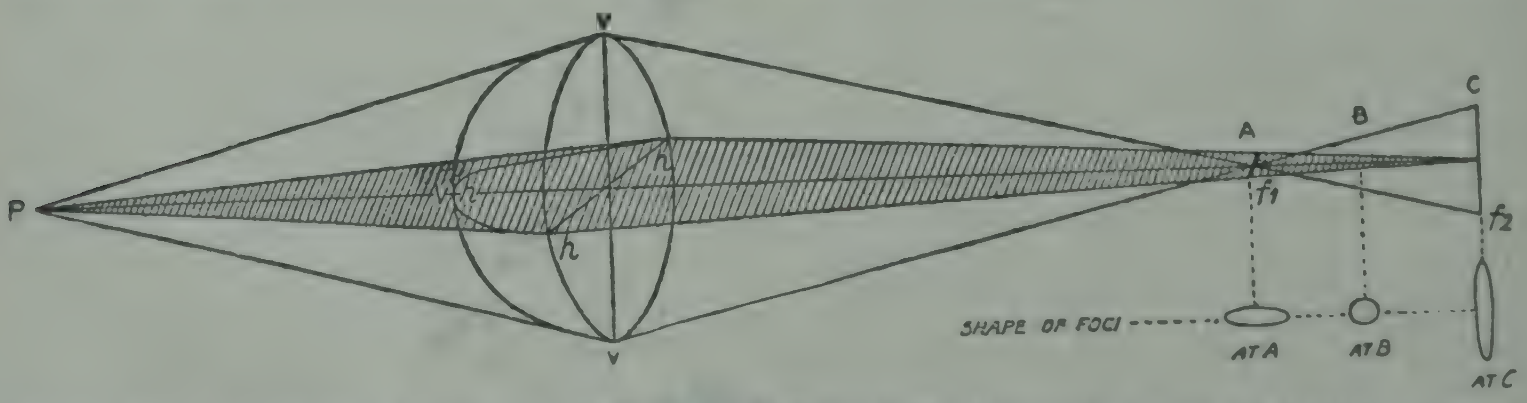


Fig. 264. Showing the Shape of Foci at different Positions in cases of Astigmatism.

In the second variety, or regular astigmatism, the parts of any one meridian give the same focus, but the different meridians have different foci (Fig. 264). There are, however, two meridians at right angles to one another, one of which has the longest and the other the shortest focus, the meridians in between showing an orderly sequence between these two extreme values. Thus the use of the term "regular astigmatism." Two types of patient are found to suffer from this condition, those who have inherited and those who have acquired it as a sequence to injury, operation or disease. The effects on vision are varied, but the characteristic features are distortion of objects looked at, and indistinctness of lines in one direction, while those at right angles are quite sharp. Headaches, eye-strain and dimness of vision are very common. Many types are met with because the maximum and minimum meridians may occupy any angle, so long as they are at right angles to one another, and there may be any accompanying degree of myopia or hypermetropia.

The diagnosis and measurement of astigmatism present no difficulties. Its existence may be readily proved by causing the patient to look at a figure consisting of a series of lines radiating from a common centre. It is then found that while some of the lines are sharp those at right angles are indistinct. This test also shows the axes of the principal meridians. By retinoscopy (see p. 396) the axes and the amounts of the abnormality in those axes may be readily determined. The treatment consists in giving spectacles which have been ground on one side to a cylindrical surface. The axis of this cylinder is adjusted to correspond with one of the principal meridians of the eye of the patient. The curve given to the cylinder is that which will cause the focus of the meridian with which it corresponds to be equal to the focus of the other meridian. The other side of the spectacle lens is ground to that spherical surface which will make the eye emmetropic after it has been corrected by the cylinder.

ANISOMETROPIA. The last abnormality of refraction which we have to consider is called anisometropia. It simply means difference between the refraction of the two eyes. The effect on vision is very slight, since it is found that as a rule one eye does all the work, and the image of the other, which is necessarily indistinct, is prevented from reaching consciousness. The result in course of time is that the unused eye loses to a considerable extent its power of seeing and as a result strabismus develops. Treatment consists in giving glasses which correct each eye separately, and then instituting exercises for the poorer eye, in order to improve its vision. The results of this treatment are good.

X-RAY MEASUREMENTS ON THE LIVING HUMAN EYE. The X-ray methods have been used recently in order to make two kinds of measurements on the living human eye: (*a*) The dimensions of the globe of the eye and (*b*) the focal length of the lens system of the eye. The methods depend on two facts, that the incidence of X-rays on the retina produce a sharply localised visual effect, and that these rays are not deflected during their transmission by the tissues. Prior to the employment of these methods it was only possible to measure the diameters of the globe of the eye after death and the focal length of the lens system of the eye was evaluated by complicated optical methods, tiring to the subject and exacting to the observer. Now, both classes of measurement can be made on living human subjects quickly and easily.

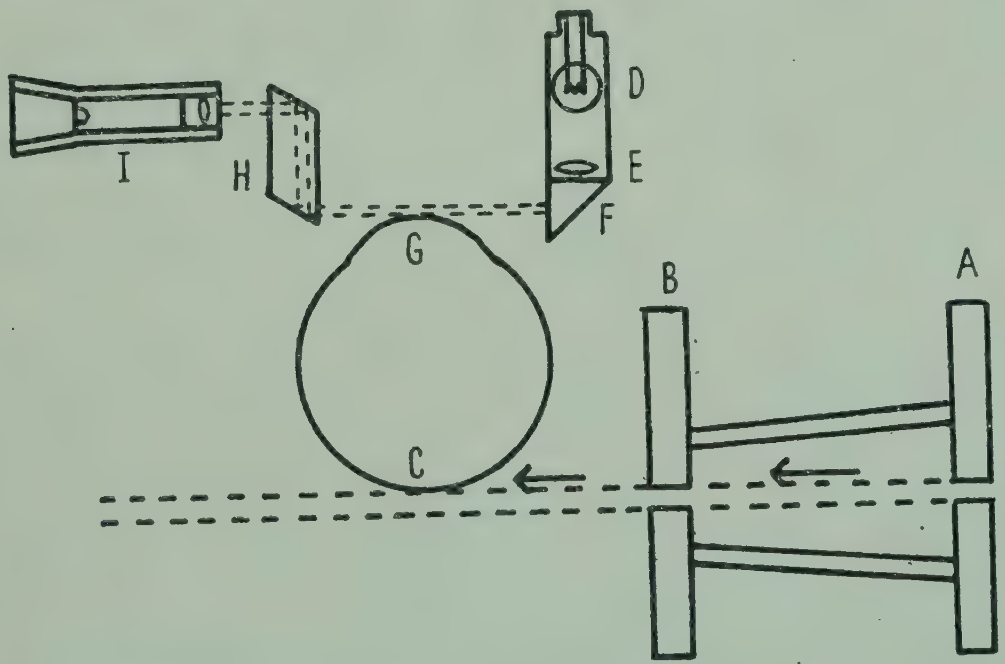


FIG. 265. The X-ray measurement of the axial length of the living human eye.

The X-ray apparatus AB is adjusted until its beam just traverses the retina at C. Light from the lamp D rendered parallel by the lens E is reflected by the prism F past the cornea and is observed by the operator through the prism H and the low power microscope I. When both X-ray apparatus and microscope apparatus have been correctly adjusted the axial length of the subject's eyeball is indicated by the distance between these two beams. (After RUSHTON, R. H., 1938, *Trans. Ophthal. Soc. U.K.*, **58,** 136.)

In order to measure the front to back dimensions of the eye Rushton used the methods illustrated diagrammatically in Fig. 265. A narrow ribbon of X-rays is caused to graze the posterior surface of the retina by adjusting the position of the apparatus relative to the patient until he sees in the dark a tiny bright disc of light. By means of a low power micrometer microscope the position of the summit of the cornea is now determined relative to the X-ray apparatus, and thus the antero-posterior length of the eye is measured.

The side to side and top to bottom dimensions are ascertained by first causing the X-ray beam to graze the retinal surface on one side and then on the other. Some measurements made by Sorsby and O'Connor using this method are shown in Table I. Myopic eyes are usually found to have greater lengths and hypermetropic eyes shorter lengths than those of emmetropes. But this is not always the case and the latter are found to vary considerably among themselves.

TABLE I

THE DIMENSIONS OF THE LIVING HUMAN EYE (SORSBY AND O'CONNOR)

Case.	X-ray measurements (mm.).			Refraction (Dioptres).	
	Transverse.	Vertical.	Axial.	Vertical.	Horizontal.
1	21·5	20·5	22·0	0	0
2	23·5	23·0	23·0	+ 1	+ 2
3	24·0	24·5	25·0	− 3·5	+ 2
4	23·5	24·0	23·0	+ 1·0	+ 1·5
5	24·5	25·0	28·0	− 16·0	− 14·5
6	24·5	23·5	24·5	0	+ 0·5

Goldmann and Hagen have used the X-ray method for the second purpose mentioned above, namely for ascertaining its focal length, or, to be precise, its posterior nodal distance. For this purpose two parallel ribbon-shaped beams of X-rays are directed towards the posterior part of the retina so that they lie one on either side of the fovea (see Fig. 266). The consequence is that the subject sees two bright lines projected into space in front of him. Two movable bright pointers are now adjusted in turn until their tips correspond to these bright lines. Now the distance apart of (*a*) the X-ray beams, and (*b*) the pointers, and (*c*) the distances of the pointers from the subject being known, the posterior nodal distance "N" may be calculated from similar triangles by means of the formula :

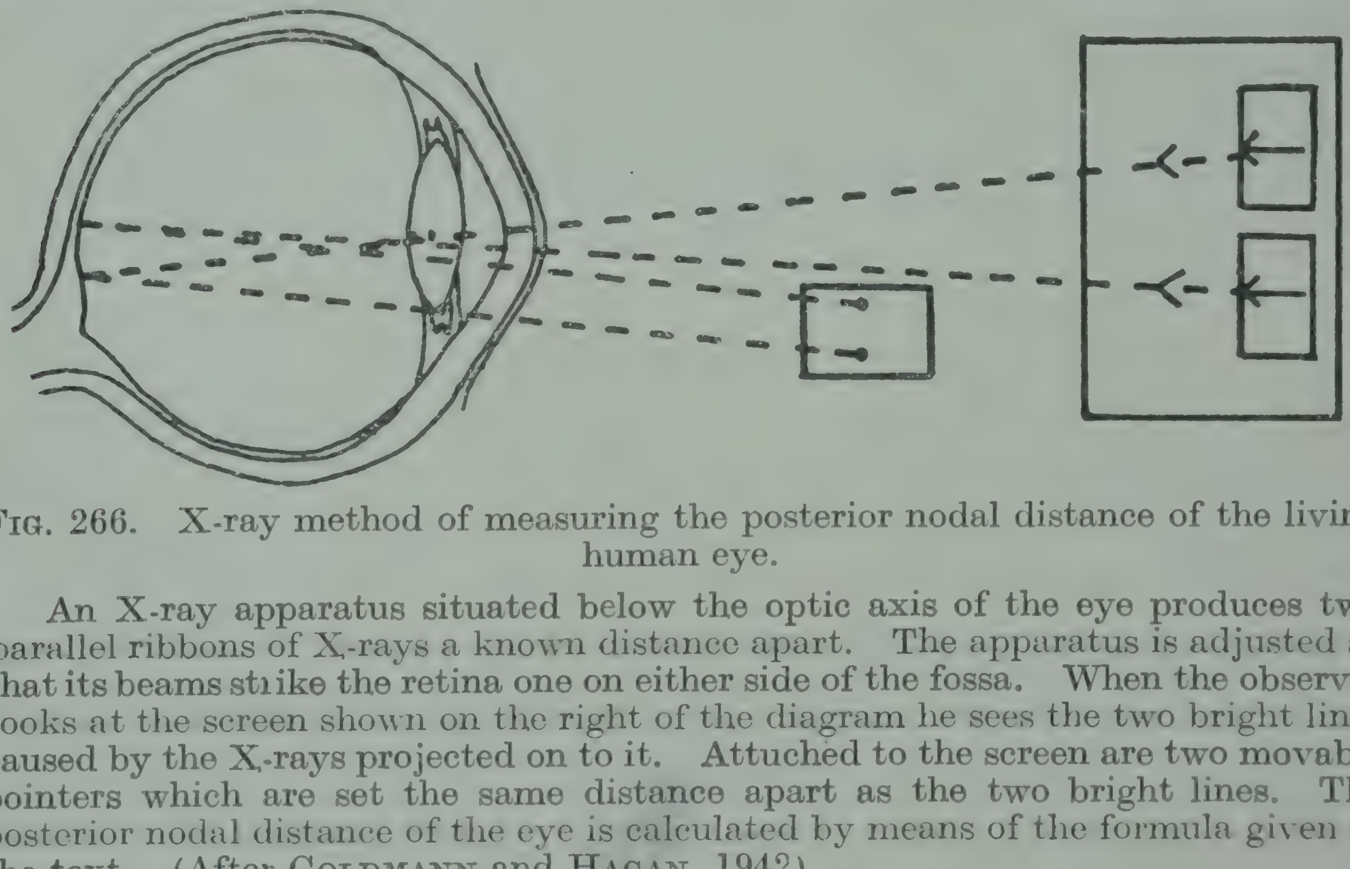


FIG. 266. X-ray method of measuring the posterior nodal distance of the living human eye.

An X-ray apparatus situated below the optic axis of the eye produces two parallel ribbons of X-rays a known distance apart. The apparatus is adjusted so that its beams strike the retina one on either side of the fossa. When the observer looks at the screen shown on the right of the diagram he sees the two bright lines caused by the X-rays projected on to it. Attached to the screen are two movable pointers which are set the same distance apart as the two bright lines. The posterior nodal distance of the eye is calculated by means of the formula given in the text. (After GOLDMANN and HAGAN, 1942).

$$N = \frac{c}{b} \times a$$

Some values obtained by Goldmann and Hagen are shown in Table III.

THE HYPOTHESES OF ABNORMAL REFRACTION. There have been three main hypotheses concerning the causes of the errors of refraction : (1) That they are largely due to errors in the length of the eyeball, a myope having too long an eyeball, and a hypermetrope having too short an eyeball ; (2) that they are caused by the chance association of two completely independent variables : the length of the eyeball and the posterior nodal distance of the lens system of the eye ; if this were true, then refractive errors would be found to fall into nine classes, as shown in Table II. (3) That emmetropia is more common than either of the above hypotheses would suggest, owing to eyeball length and posterior nodal distance undergoing similar variations. Thus an emmetrope with an eyeball of normal length has a lens system with an average posterior nodal distance, and one with a long eyeball has a long posterior nodal distance, and *vice versa*. If this hypothesis be true there must be some association between the factors which cause a long eyeball and those which cause a long posterior nodal distance, and *vice versa*.

TABLE II

Eyeball length.		Long.	Medium.	Short.
Posterior Nodal Distance	Small	High myope.	Myope.	Emmetrope.
	Medium	Myope.	Emmetrope.	Hypermetrope.
	Large	Emmetrope.	Hypermetrope.	High hypermetrope.

Which of these hypotheses is the correct one ? In Table III are given some of the results obtained by Goldmann and Hagen by means of their X-ray method. The last column shows that the posterior nodal distance may vary from 14·9 to 17·83 mm. ; and therefore the first hypothesis must be incorrect. Dealing similarly with the second hypothesis, if there is completely independent variation of eyeball length and posterior nodal distance, one would expect the average values for the latter would be the same for hypermetropes, for emmetropes and for myopes. Now the average for 10 emmetropes was found to be 16·81 mm., that for 6 myopes was 17·04 mm., and that for the single hypermetrope was 14·90 mm. Now it would be greatly to the advantage of this analysis for there to be averages for many more subjects. But so far as the figures available go, the averages of the three classes vary too much for the second hypothesis to be correct. That the third hypothesis is probably correct is shown by an examination of the posterior nodal distances for the emmetropes, which varied from 15·38 mm. to 16·69 mm. ; in order that these subjects should be emmetropic it would be necessary for compensatory changes in eyeball length to occur. This idea is supported by an examination of the posterior nodal distances of the hypermetropes and the myopes, for they would have been more seriously hypermetropic and myopic, respectively, if it had not been for the existence of compensatory variations in their posterior nodal distances.

TABLE III

GOLDMANN AND HAGEN'S DATA

Subject's number.	Class of refraction.	Correction in dioptres.	Total refraction in dioptres.	Posterior nodal distance mm.
1	Emmetrope	—	61·0	16·39
2	,,	—	57·9	17·27
3	,,	—	62·1	16·10
4	,,	—	57·9	17·27
5	,,	—	59·0	16·95
6	,,	—	65·0	15·38
7	,,	—	60·2	16·61
8	,,	—	58·7	17·04
9	,,	—	57·0	17·54
10	,,	—	56·6	17·67
11	Hypermetrope	+ 1·0	67·1	14·90
12	Myope	− 1·75	58·6	17·06
13	,,	− 3·5	61·4	16·29
14	,,	− 2·25	56·1	17·83
15	,,	− 2·0	57·1	17·51
16	,,	− 2·5	62·4	16·03
17	,,	− 4·25	58·1	17·21
18	,,	− 5·0	57·2	17·48

6. THE RETINA

The retina is a delicate membrane lying inside the choroid coat of the eye. Its internal surface lies in contact with the hyaloid membrane of the vitreous body. It is thus supported on both sides. The retina itself consists of two layers, the outer or pigmented, and the inner or nervous. Whereas embryologically the retina covers the whole internal surface of the eye including the ciliary processes and the iris, this is not the case with the nervous layer, because this stops near the equator of the eye at the ora serrata, and is here replaced by a layer of columnar epithelium. Opposite the pupil a yellow spot is seen on the retina, the *macula lutea*, and in the centre of this there is an oval depression, the ***fovea centralis***. The optic nerve enters the eyeball through an aperture in the sclera and choroid, and then passes through the posterior surface of the retina to spread out over the internal surface. In the fovea, however, this is not the case, for the depression at this point is caused by the absence of nerve fibres. The point at which the optic nerve enters the eye is easily recognised from inside the eyeball because the numerous white nerve fibres, as they bend over the edge of the aperture in the retina, form a characteristic white mound called the colliculus, at the centre of which is a depressed portion called the optic cup. It is in the centre of this cup that the central artery of the retina and the corresponding vein first make their appearance. These have the important function of nourishing the retina; the additional blood supply through the intimate contact between the retina and the vascular choroid, although important, is quite insufficient to supply the needs of vision, as is shown by the immediate and permanent blindness which follows blocking of the central artery of the retina.

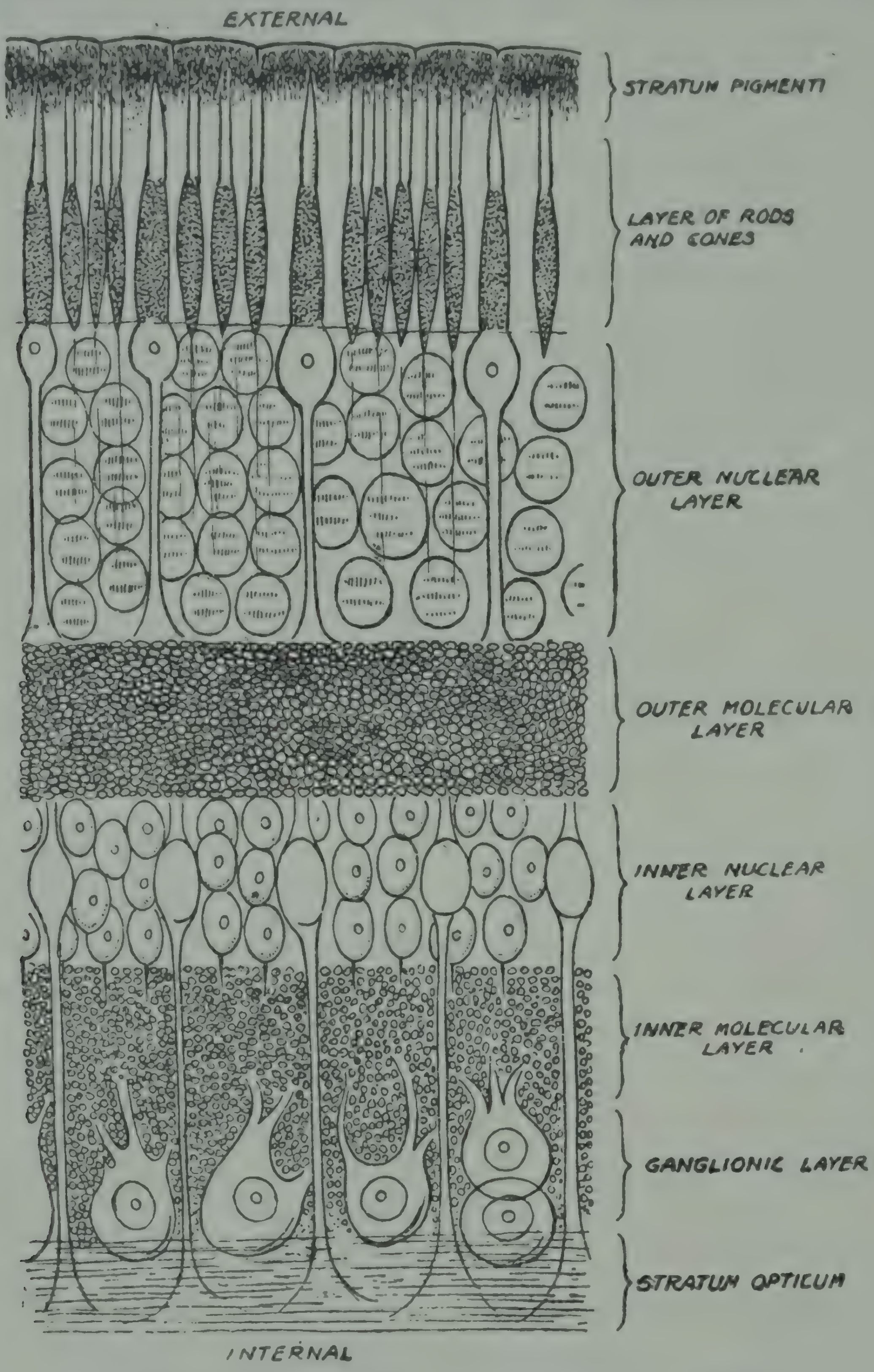


Fig. 267. Diagram of Transverse Section of Retina.

When sections of the retina are examined under the microscope it is found that they consist of the following layers from without inwards (Figs. 267 and 268):—

1. Layer of pigmented epithelium	Developed from posterior layer.
2. Layer of rods and cones (bacillary layer)	Developed from anterior layer of optic vesicle.
3. Outer nuclear or granular layer	
4. Outer molecular or plexiform layer	
5. Inner nuclear or granular layer	
6. Inner molecular or plexiform layer	
7. Layer of ganglion nerve cells	
8. Layer of nerve fibres and vessels	

In order to understand the structure of these layers, it is necessary to keep in mind the fact that the optic nerve and cup are outgrowths of part of the brain. We must therefore be prepared to find in the retina all those structures which are found in every sensory path to intervene between the sense cell and the brain nucleus. In every case the stimulus is conveyed through three sets of neurones or relays before it reaches the brain.

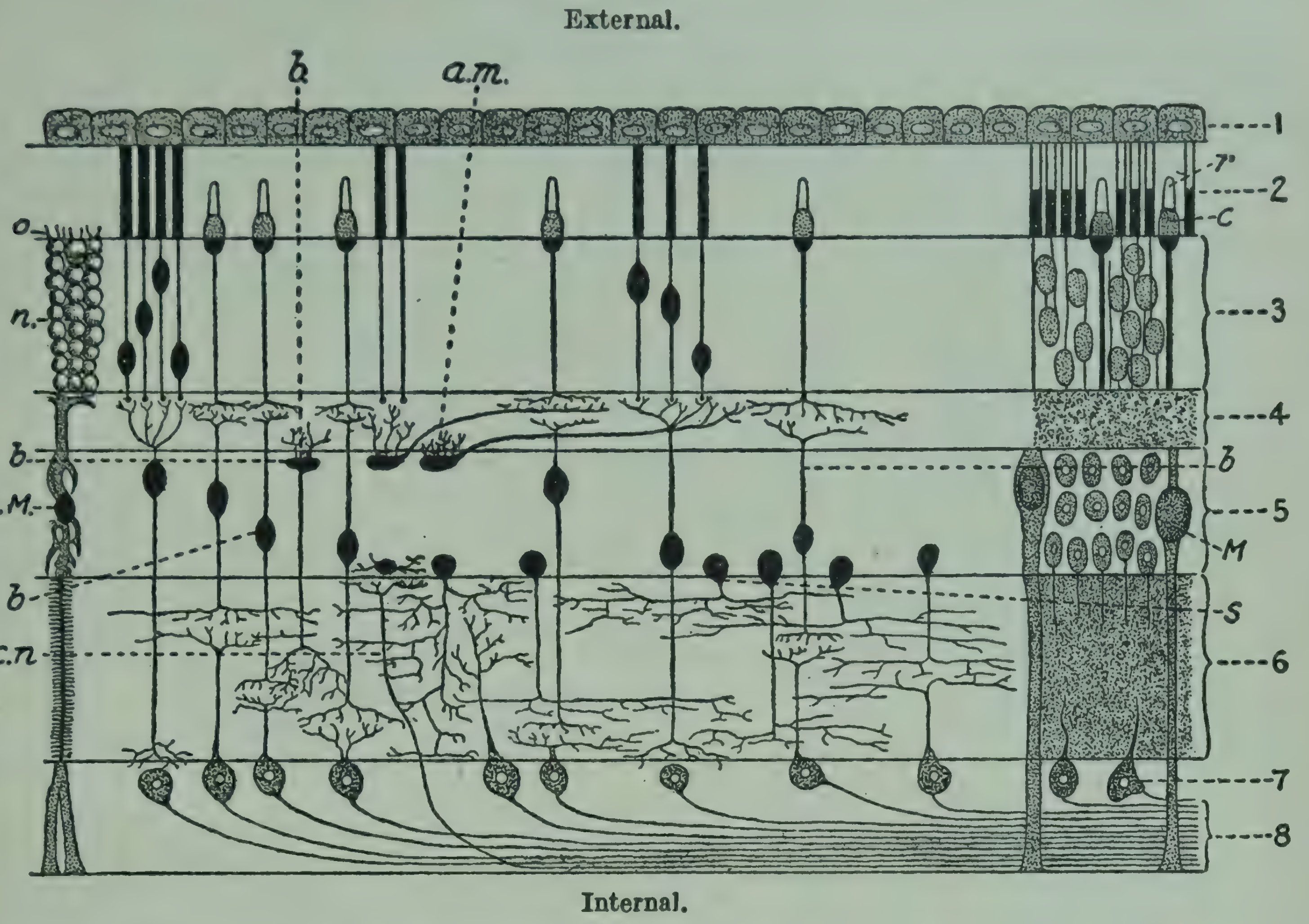


Fig. 268. Diagram to show the different Cellular Structures found in the Retina.

The Stratum Pigmenti is the only one that is developed from the external layer of the embryonic optic cup. The epithelium consists of a single layer of hexagonal nucleated cells containing numerous pigment granules. The cells send fine processes between the limbs of the rods. The bases of these cells are firmly attached to the choroid, and thus give support to the rest of the retina.

The object of these cell processes and the pigment granules which they contain would appear to be either the prevention of an image formed on one part of the retina from spreading to the sensitive elements of surrounding portions, or else the protection of these elements from excessive light action. But it has been definitely proved that the cells themselves have another and important function to perform, namely the secretion of the pigment called visual purple (rhodopsin). The important functions of this pigment will be described later.

The Bacillary Layer of rods and cones is separated from the succeeding layer by the external limiting membrane. Each rod and cone consists of an outer and inner limb, the forms of which are well shown in Fig. 269. It will be seen that the outer limbs are striated, the cones coarsely, the rods finely; they tend after hardening to break up into discs. The inner limbs of both rods and cones have a strong affinity for dyes.

The Outer Nuclear Layer consists of the cells of the first order neurones or the granules of the rods and cones. The cells are nucleated, somewhat smaller than the bipolar cells, and their nuclei are striated. They give off two processes, one of which forms dendrites in the fourth layer, the other connects with either a rod or cone as the case may be (see later).

The Outer Molecular Layer is much like the inner; it consists of the dendrites of the second order neurones and the first.

The Inner Nuclear Layer largely consists of bipolar second order neurone cells. There are, however, also present the nuclei of the horizontal cells, and also the nuclei of similar cells (amacrine cells, *a.m.*, Fig. 268), whose dendrites travel in the outer molecular layer. The bipolar cells, which are fusiform in shape and nucleated, are of three kinds: (*a*) those which connect with rods, (*b*) those which connect with cones, and (*c*) giant bipolars which connect with either.

The Inner Molecular Layer consists of a feltwork made up by the interlacing dendrites of the ganglion cells with those of the inner nuclear cells or second order neurones. There are also the dendrites of horizontal cells or spongioblasts (*s*, Fig. 268). These possibly serve to associate the impulses from different parts of the retina, such as is supposed to occur in the brain. It should be noted that they appear to be absent from the fovea and macula.

The Ganglion Cell Layer consists of a single layer of large oval cells. These are nucleated and give off the axons which we have already described, and a bunch of dendrites which ramify with others in the inner molecular layer. Only at the macula is more than one layer of ganglion cells present; this is due to their almost complete absence at the fovea. The macula therefore not only has its own relays but those of the fovea as well.

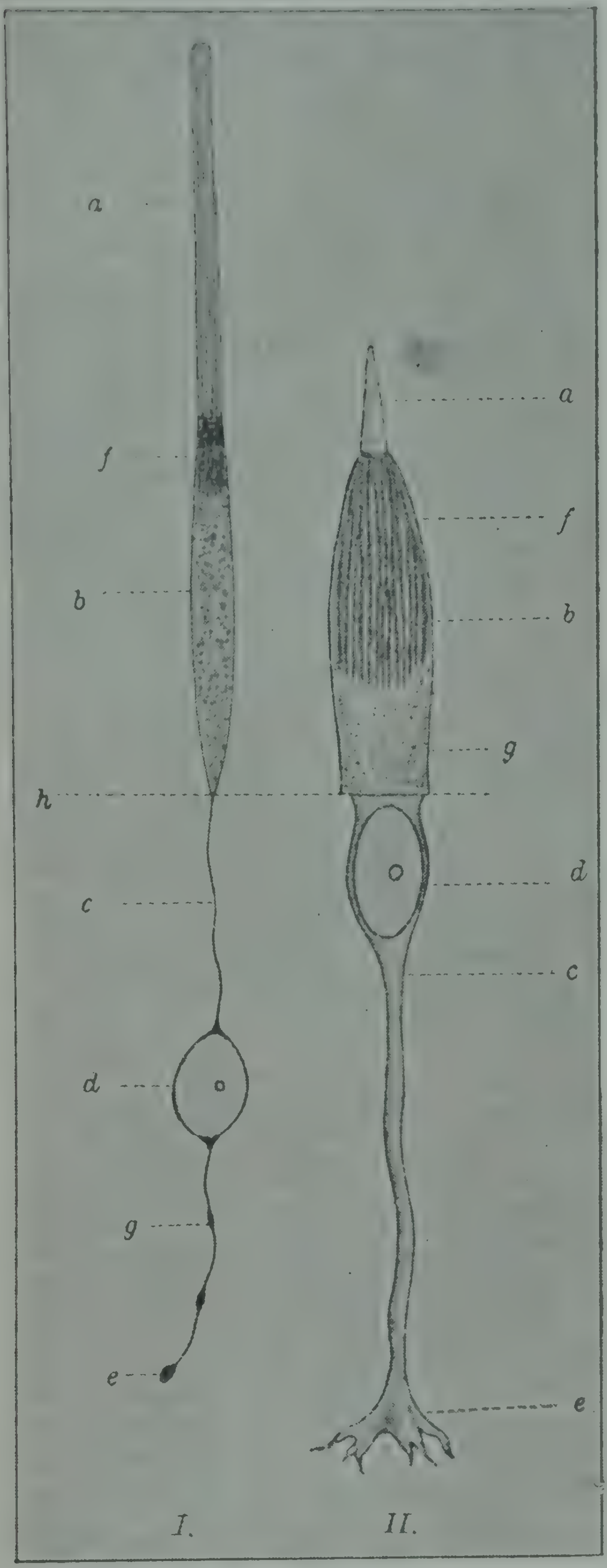


Fig. 269. (R. Greeff.)

I. A Rod.
II. A Cone of Mammalian Retina.
h. External Limiting Membrane.

The Nerve Fibre Layer (stratum opticum) consists of the non-medullated axons of the large ganglion cells found in the second layer. These axons are the third order neurones which become myelinated after they have passed out of the eyeball, and travel by paths to the calcarine cortex described on p. 287. Beside the fibres conveying visual impressions there are others which belong to the pupillomotor reflex. Others again bring impulses from the brain to the retina; their functions will be considered below.

It should be noted that beside the structures described above, which have the functional activities of the retina to perform, there are a number of connective tissue elements which form the retina into one coherent structure. Since the retina is developed from an out-growth of the brain, these structures are similar in type to those met there; we therefore find neuroglia and also long cells which extend through the first seven layers and hold them together, namely the fibres of Müller (*M.* and *n.M.*, Fig. 268).

The Different Parts of the Retina show marked variation in detail. At the fovea (Fig. 270) cones alone are found; only one of these connects to each axon. The structural peculiarities of the fovea may be summarised as follows:—(1) Absence of rods. (2) The cones are longer, more highly developed, and some say more rod-like than those found elsewhere. They are very closely packed, so that their inner limbs are seen in transverse section to have a hexagonal shape, the flat surfaces being in contact with those of their neighbours. (3) The rows of nerve cells and dendrites, which in the rest of the retina are approximately in line with the rod or cone to which they belong, are in the fovea pressed to one side in a direction away from the centre. In this way a cone may have the nerve cells to which it is connected placed at a considerable distance away in the surrounding macula. It is this displacement of the nerve fibres and their cells that causes the fovea to appear hollow. The purpose of this physiological arrangement would without question appear to be the avoidance, at this important region of the retina, of the scattering of the image which passage through the nerve cell layers would introduce. (4) The fovea, unlike the rest of the retina, is devoid of blood vessels. The purpose of this arrangement would appear to be similar to that just described. (5) Visual purple is said to be absent from the fovea. This would appear to be connected with the absence of rods.

Round the fovea is a ring in which rods and cones are present in almost equal number. In still more peripheral regions cones are relatively few, and several rods connect with each axon: this reduces the relative number of nerve fibres.

CHANGES IN THE RETINA ON EXPOSURE TO LIGHT

A light stimulus falling on the retina causes a number of changes to occur which may be classed as structural, physical, chemical and physiological.

Structural Changes occur on exposure of the eye to light: firstly, movement of the pigment from the outer epithelial layer into the spaces between the rods and cones; secondly, shortening of the cones themselves (Fig. 271). These changes occur only when the connections of the eye with the brain are intact. The rate of movement appears to vary with intensity, and violet

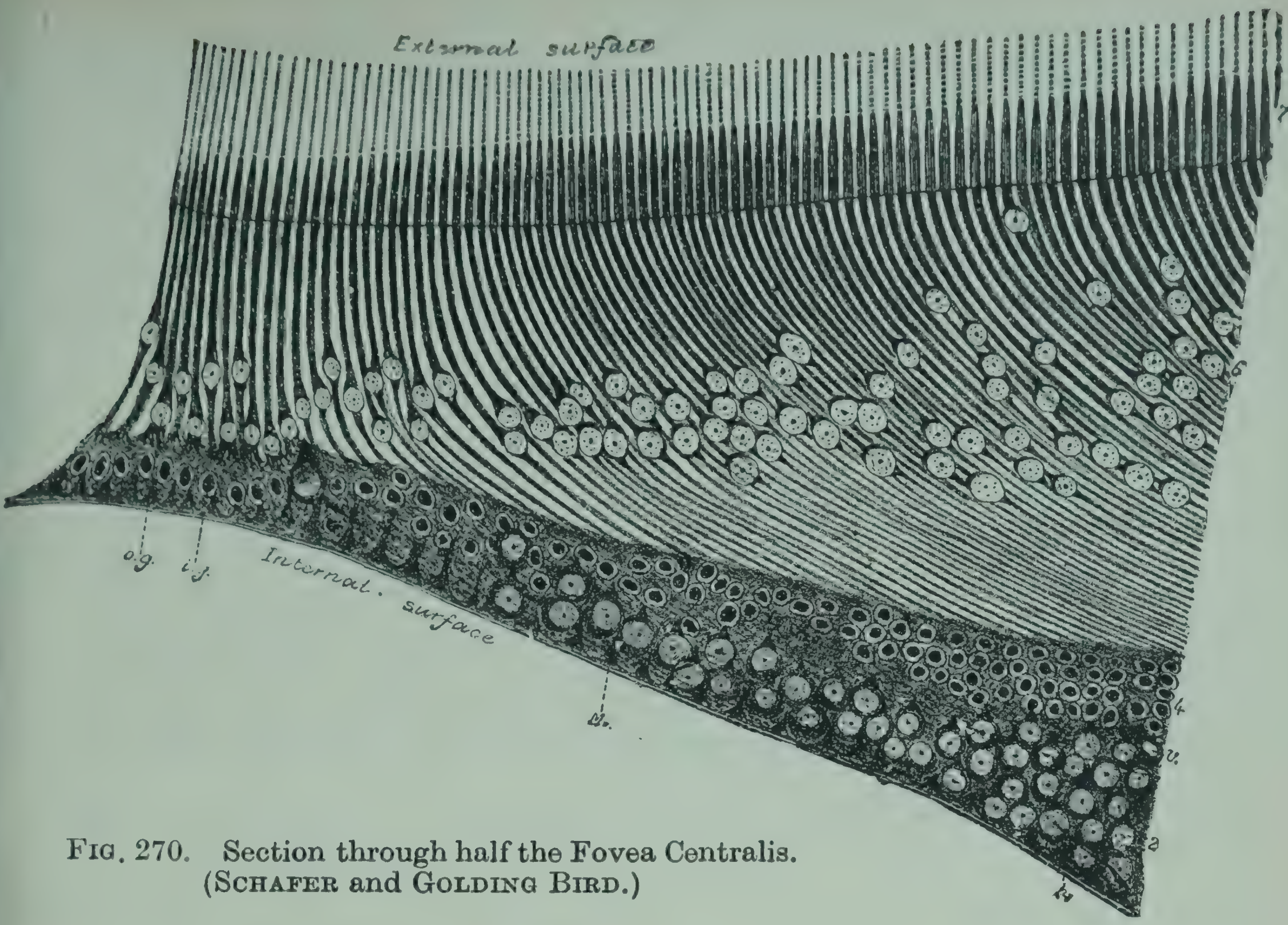


FIG. 270. Section through half the Fovea Centralis. (SCHAFER and GOLDING BIRD.)

light is said to be more effective than red. It is interesting to find that electrical stimulation of the optic nerve or the falling of light on the other retina to that of the eye observed also causes these cone movements. It is supposed that the impulses which effect these movements travel through the nerve fibres already described as descending from the brain to the retinæ; it is for this reason that Engelmann called these fibres "retinomotor."

Other structural changes that are found, by histological investigation, to follow exposure of the retina to light, are swelling of the outer limbs of the rods, and the disappearance of chromatin granules from the ganglion cells. Both these changes are said to occur more rapidly under the action of rays of short wavelength.

PHYSICAL CHANGES are also found when the retina is stimulated by light, namely an electrical response somewhat similar to the current of action in

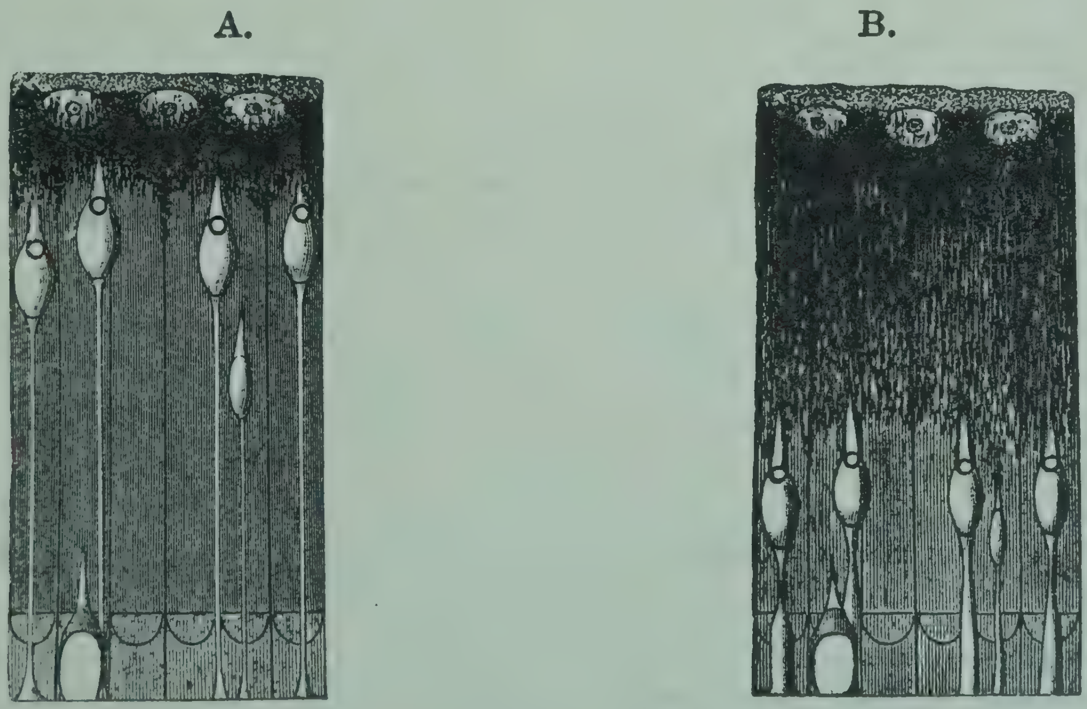


FIG. 271. Sections of the Frog's Retina.

A. Kept in the dark. B. After exposure to the light, showing retraction of the cones, and protrusion of the pigmented epithelium between the outer limbs of the rods. (ENGELMANN.)

nerve. A typical curve is shown in Fig. 272. A study of the electric response to lights of different colours and intensities has given the following results. With light of any one colour, a geometric rise of intensity causes an arithmetic increase in the current. With coloured rays of apparent equal intensity, yellow rays are said to give a larger current in the light adapted eye, and green in the dark adapted eye. It is interesting to observe that the current commences after a latent period, which is of the same order as that found for the perception of light by the eye. This and other facts mentioned above, would seem to point to the currents observed being the accompaniment of the passage of the nervous impulses to the brain.

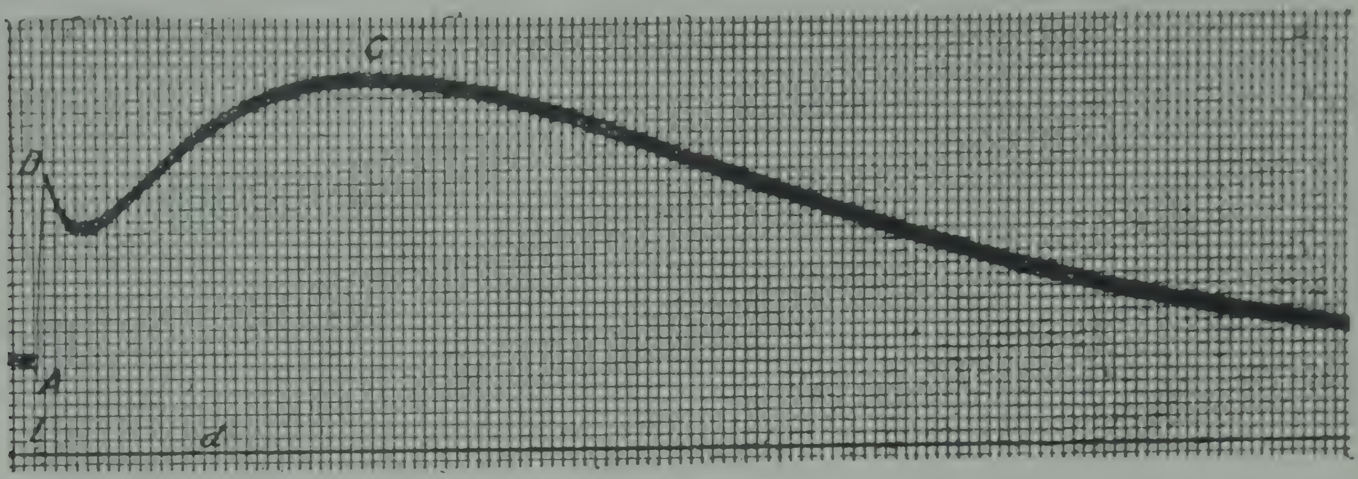


Fig. 272. Electrical Variation in Frog's Eye as recorded by the String Galvanometer. (Einthoven and Jolly.) On exposure to a single flash.

Chemical Changes in the retina on exposure to light are of two kinds: firstly a tendency of the retina as a whole to become acid in reaction, as shown in the change in its behaviour to certain stains; and secondly the bleaching of two pigments, namely the *visual purple* and *fuscin*. With regard to visual purple (or rhodopsin) a large number of facts have been made out. In the first place it is found in association with the rod retinal structures only, and is therefore absent from the human fovea centralis. It is a derivative of vitamin A, and when bleached by light liberates various products related to the carotenes.* Transient orange first appears and then indicator yellow (Lythgoe).† Bleaching on exposure to light occurs in solution as well as in the retina. According to Wald exposure to light causes visual purple to split into a yellow pigment "retinene" and a protein. Retinene on further exposure to light yields a colourless substance—visual white. When vitamin A is present in the retina the protein can combine with it to form visual purple.

The absorption of light by visual purple is greatest for the middle of the spectrum; the transmitted colour is therefore composed principally of red and violet rays.‡ Small variations of colour are found, however, in samples obtained from different animals; some are even rose-pink in colour. In the second place the most potent colours in producing bleaching are those near the centre of the spectrum, that is the rays which are most strongly absorbed. This pigment therefore obeys Draper's law, which states that those rays that are absorbed are those which produce chemical change.

It will be shown below that the retina when adapted for vision in dim light is not only colour blind, but also that the green rays of the spectrum appear to have the maximum luminosity. It will also be shown that the rod-containing parts of the retina, where visual purple is to be found, are the only regions that react to light of low intensity. Moreover, the curve which represents the rate of bleaching of visual purple by rays of different colour is very similar in shape to that which shows the visibility of lights of low intensity, as in Fig. 273. The obvious inference is that visual purple is concerned with twilight vision. Accordingly it is found that animals deprived of vitamin A, a precursor of visual purple, show deficient twilight vision (night-blindness).* When the visual purple in the retina has been bleached by exposure to light, there follows a gradual re-formation of purple, which is independent of nerve connections, but occurs only so long as the stratum pigmenti is in contact with the rod epithelial layer. If we suppose that the product formed by the bleaching of the visual purple stimulates the rod apparatus, causing it to send impulses to the brain, we have a plausible hypothesis of the mechanism for night vision. This we may briefly describe as follows :—When light falls on the retina, certain rays, particularly those near the middle of the spectrum, are absorbed by the visual purple. The pigment is bleached in proportion to the light absorbed, forming a new product : this acts on the rods, causing them to send impulses to the brain which continue so long as the light falls. When the light stops the stimulating product disappears again and therefore the stimulus to the rods ceases.

---

* Wald. *J. Gen. Physiol.*, 1935, **19**, 351.

† Lythgoe and Quilliam. *J. Physiol.*, 1938, **94**, 399.

‡ Lythgoe. *J. Physiol.*, 1937, **89**, 331.

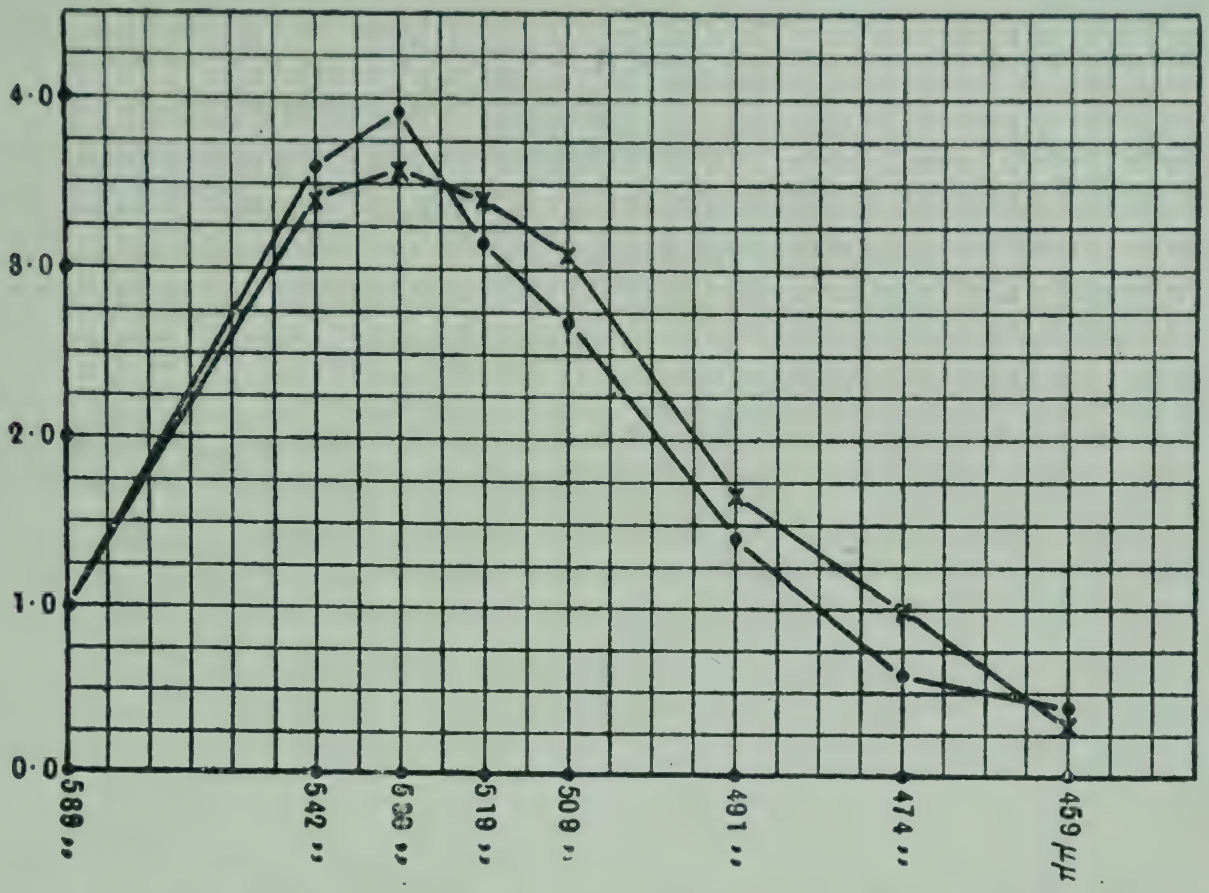


FIG. 273. Shows the Similarity between the Curve representing the Rate of Bleaching of Visual Purple by Light of different Wavelengths and the Luminosity Curve of Twilight Vision.

The work of Hecht (1920–21) is very strongly in favour of such a view. He has shown that the rate of bleaching of a solution of visual purple by light and the rate of light adaptation of the human eye both follow a logarithmic law : that is, the "reaction" in both cases appears to be one involving the splitting up of a single large molecule into components. On the other hand, the rate of dark adaptation of the human eye follows a curve similar in shape to that representing the rate of regeneration of visual purple.† Taken in conjunction these pieces of evidence point most strongly to the conclusion that a single large molecule of visual purple splits under the action of light into components, and that these when the light is removed combine again to re-form visual purple. Many analogous photochemical systems have been studied.

A picture, formed by the bleaching of the visual purple in those parts of the retina which correspond to the high lights of the image formed by the lens, can be fixed, much like a photograph, by immersing the retina in a solution of alum.

Fuscin is the pigment found in the form of needles, plates or prisms in the processes of the cells of the stratum pigmenti of the retina. The object of this pigment is apparently to absorb light, which might tend to spread from those retinal elements on which an image of a light source is falling to neighbouring ones. Some of this pigment is bleached by strong light, but so far as is known the break-down products have no visual function. The presence of other pigments has been described in the retina, such as visual yellow and the bright pigment granules found in birds. The presence of the latter is too variable for them to be considered to take any essential part in the visual mechanism.

---

* TANSLEY. *J. Physiol.*, 1931, **71**, 442.

† TANSLEY. *J. Physiol.*, 1931, **71**, 442 ; *Proc. Roy. Soc.*, 1933, **114B**, 79.

Physiological changes produced by light depend greatly on the region of the retina on which it falls, since this may contain rods only, cones only, both rods and cones, or neither rods nor cones. The peripheral parts of the retina contain numerous rods and very few cones. When stimulated by light of low intensity, this part of the retina is found to be exceedingly sensitive, particularly if the eyes have been closed or kept in the dark for a time. Tests with light of low intensity and of different colours show that the region is colour blind, but that rays in the middle of the spectrum are more readily appreciated than others. We have here well developed the so-called "twilight vision," which is associated with the rod-visual purple mechanism just described. Besides being very sensitive to light of low intensity, the periphery of the eye is also particularly perceptive of light of short duration. This part of the retina, therefore, appreciates movement at night very readily. Lastly, owing to the fact that a number of rods connect with one nerve fibre conveying the impulses to the brain, the periphery of the retina has a poor perception of detail.

THE FOVEA CENTRALIS is found to contain cones only. The vision in this region is therefore the antithesis of that found at the periphery. The appreciation of light of low intensity is bad, but when an image is sufficiently bright to cause stimulation, its colour is perceived. When light is poor, rapid motion is not so well observed as it is by the periphery. There is an extraordinary acuteness in perceiving fine detail. This is due to the fact that the cones in the fovea are very closely packed, so closely that they become flattened where they touch one another and thus have a hexagonal shape in transverse section. Further, each cone is connected to its own nerve fibre, so that no cyphering of the impulses can occur on the way to the brain. Experiments on visual acuity definitely show that the fineness of the detail, which the eye can perceive at the foveal region, is fully as great as that which we should expect to find, if each one acted quite independently of its neighbours. Parts of the retina around the macula lutea, since they contain both rods and cones, possess, as we should expect, both the power to perceive colour found at the fovea and the ability to react to light of low intensity, without colour vision, which is possessed by the periphery of the retina. The presence of rods scattered between the cones, however, naturally impairs to some extent the appreciation of fine detail. Rays of light which reach the fovea centralis by passing through the centre of the pupil were found by Stiles and Crawford to be more effective than rays which have passed through the periphery of the pupil. This phenomenon has been called by them the directional sensitivity of the

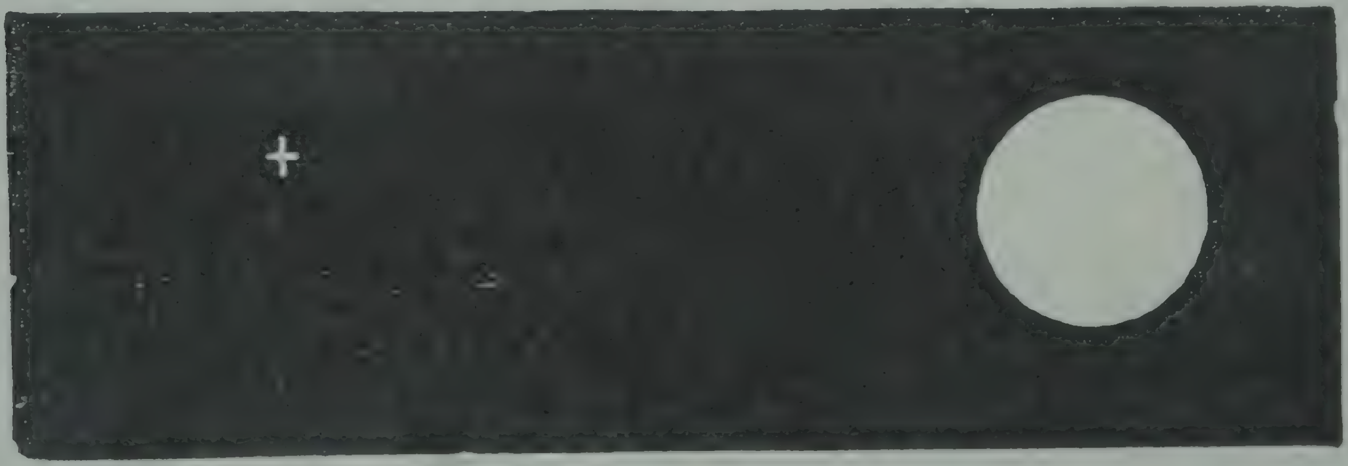


FIG. 274. Look at Cross with Right Eye, hold Book at about 10 inches.

retina. They suppose it to be due to the narrow cylindrical shape of the foveal cones (see p. 404).

At the papilla where the optic nerve enters the eye there are neither rods nor cones, and therefore, as we should expect, this region is quite blind. This fact can be readily proved by looking with the right eye at the cross in Fig. 274. If now the book be held about 10 inches from the eye the white disc will be found to disappear. By a simple calculation, it is found that the disc corresponds with the papilla of the optic nerve.

THE VISUAL FIELDS. Since the appearance of an external object will vary to a considerable extent, according to the region of the retina on which its image falls, it is a matter of considerable interest to determine the positions at which the appearances undergo change. This is also of practical value because the positions are found to be affected by disease. The determinations are usually made by means of an instrument called a *perimeter*.

This consists (Fig. 275) of a metal arm bent to the segment of a circle. This is so mounted in relationship to a horizontal bearing that the segment always has its centre in correspondence with a fixed pointer which is seen on the left of the diagram. If the eye of the patient is applied close to this pointer and looks towards the centre of the bearing, the degrees marked on the metal segment show the actual angle at which an index is situated in relationship to the eye axis, no matter in what meridian the metal segment may lie. The index mark usually consists of a disc 2 mm. in diameter, either of white or of coloured paper, according to whether the rod or cone area is to be examined.

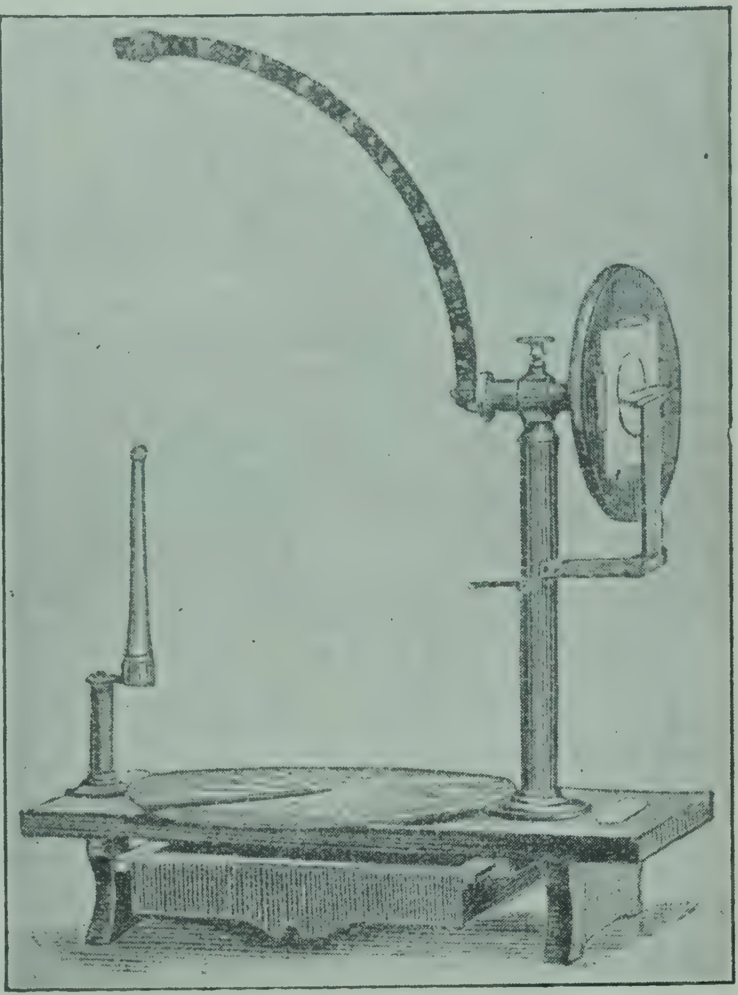


FIG. 275. Priestley Smith's Perimeter.

The results obtained by means of this instrument are shown in Fig. 276, which gives a typical curve for the right eye. The shaded area on the diagram is due to the obstruction of the eyebrow, nose and cheek of the patient. The visual field on the outer side will be seen to extend actually 14 degrees beyond the right angle. This result, which at first sight appears to be impossible, is in fact due to the considerable refraction of the light rays that occurs at the extreme edge of the cornea. The direction of the beam of light entering the eye under these circumstances is shown in Fig. 277, X.

It will be observed that, even when looking straight in front, a man can see to a considerable angle behind himself. By deviating the eyes only slightly to either side, this angle can be increased to 40 degrees in the average case, as shown at Fig. 277, Z. This ability to see to a considerable extent behind him is due to the narrowness of the head between the frontal processes and zygomatic bones. In those animals in which the eyes are placed on the sides of the head, and the visual axes are diametrically opposite to one another, the visual fields will actually overlap a short distance from the head, so that there will be no direction from which an enemy can attack without being observed (Fig. 277, Y).

UTILITY OF PERIPHERAL VISION. The high visual acuity of the fovea, and the great facility with which the eyes can be directed so that images form on this region, might raise the question as to the utility of peripheral vision. This question may be investigated experimentally by placing restricting screens in front of the eyes, or by ascertaining the experience of persons who are suffering from blindness in the periphery as the result of disease, *e.g.* retinitis pigmentosa. Both methods show that the periphery is of great value in

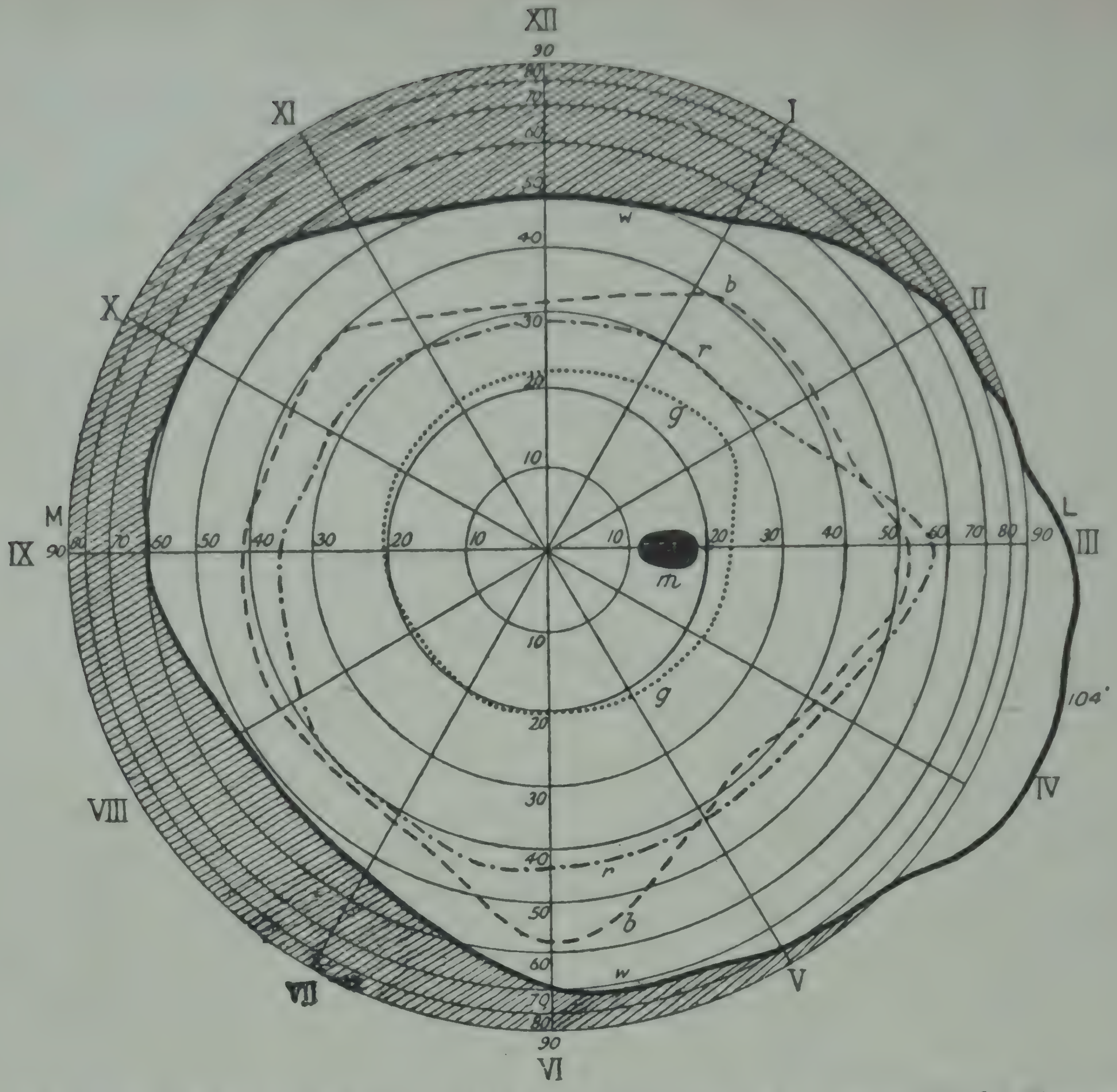


FIG. 276. Field of Vision for White and Colours of a normal Right Eye as obtained by the Perimeter. (HARTRIDGE.)

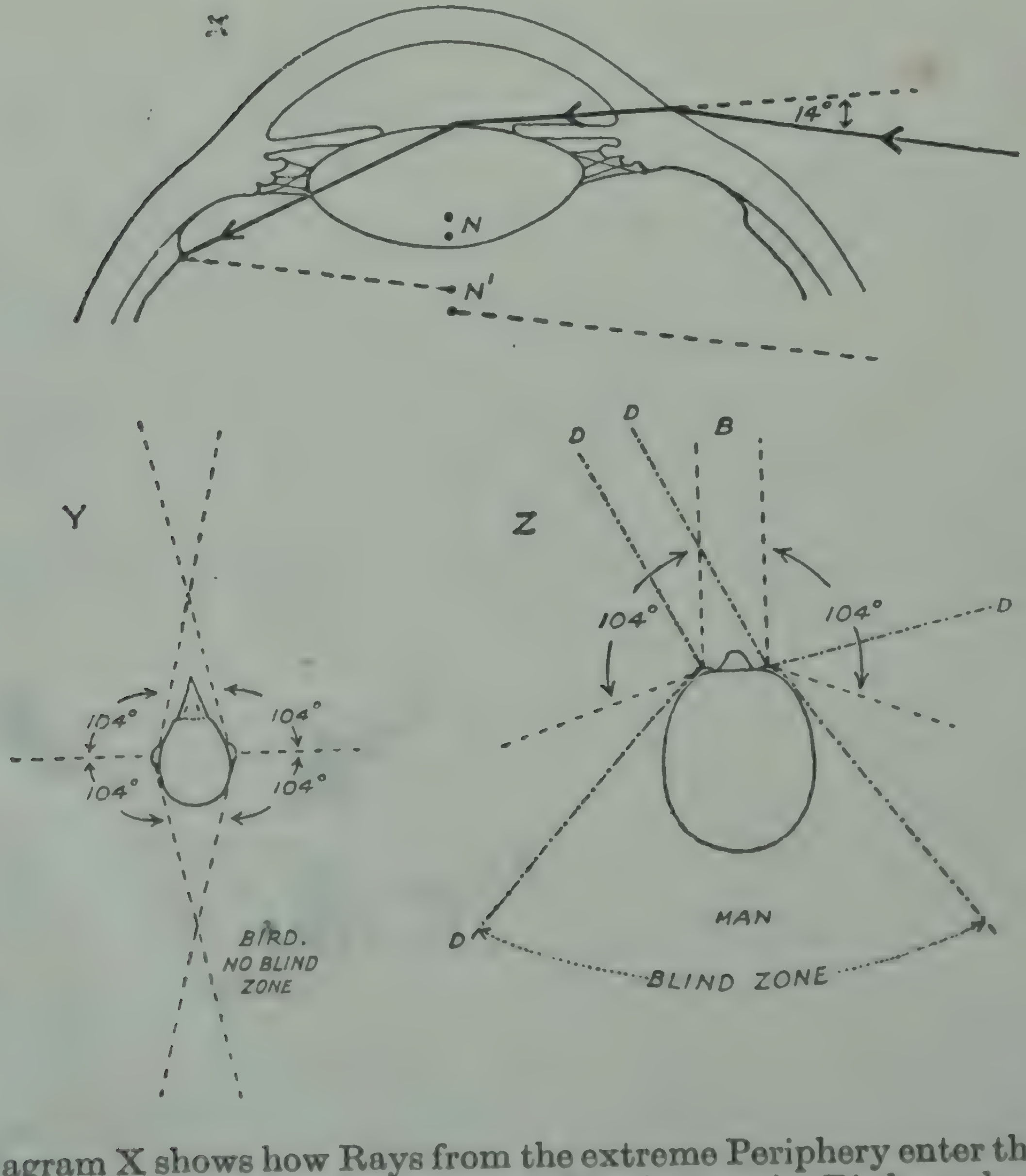

X
14°
N
N'
Y
104°
104°
104°
104°
BIRD.
NO BLIND
ZONE
Z
D
D
B
104°
104°
D
MAN
D
BLIND ZONE

FIG. 277. Diagram X shows how Rays from the extreme Periphery enter the Eye and reac the Retina. Diagram Y shows Absence of Blind Zone in Birds and certain Animals Diagram Z shows Size of Blind Zone in Man. (HARTRIDGE.)

directing attention to outlying objects. Our attention being excited, we direct the gaze in the direction indicated, in order to bring into action the greater power of analysis of the fovea.

Central Connections of the Retina. The nerve paths by which the retina, itself a part of the central nervous system, connects with other parts of the central nervous system has been considered in detail (see p. 287). Briefly the optic nerves from the two retinæ converge and join at the chiasma. These fibres from the nasal halves of both retinæ cross, so that the optic tracts which convey the impulses from the chiasma to their destinations contain fibres which have come from the same side of the retinæ as that to which they themselves belong. The right-hand tract has fibres from the right-hand halves of both retinæ. The fibres pass chiefly to the calcarine region of the cerebral cortex *viâ* the lowest part of the optic thalamus and the most posterior part of the corona radiata for the sense of sight. But some pass *viâ* the superior corpora quadrigemina to the nuclei of the oculomotor nerves, thus causing reflex eye movements.

THE OPHTHALMOSCOPE. If the retina of a patient be illuminated by causing a beam of light to enter the pupil, the reflected light will cause the interior surface of the retina to be visible. In order to see the image distinctly, it is necessary either that both the eye of the patient and also that of the observer should be focussed for infinity (the direct method) or that both eyes should be focussed for one and the same intermediate plane (the indirect method). Both methods should be practised, since the indirect method gives a low magnification while the direct method gives a high one. The indirect method is carried out as follows :—A bright source of light having been placed behind and slightly to one side of the patient, the observer standing about half a metre in front of him reflects by means of a concave mirror an image of the light into his pupil. At the centre of the mirror is an aperture through which he sees the light which is reflected back from the patient's retina. The observer now holds a biconvex lens of about 6 cm. focal length about 8 cm. in front of the patient's eye, while he still directs the beam of light into the pupil as before. The image of the retina, which would normally be focussed by the lens system of the patient's eye at infinity, is now brought to a focus by the convex lens, forming what is called an aerial image (see Fig. 278). It is this image that the observer sees (Fig. 279) and in it are shown all the particular features of the vessels and nerves of the patient's retina. Beside its very great utility to the oculist, the ophthalmoscope is a very valuable instrument to the physician, for retinal vessels and nerves frequently show the evidences of constitutional disease, which are of great assistance to diagnosis. Ophthalmoscopes are usually fitted with a number of small lenses of graduated power, which may be introduced as required behind the mirror. For the indirect method they are seldom required. The magnification of the retina given by the indirect method is usually about 3, while that provided by the direct is 12. If a higher magnification is an advantage it may be obtained by the above method by using a biconvex lens of longer focal length (say 12 cm.).

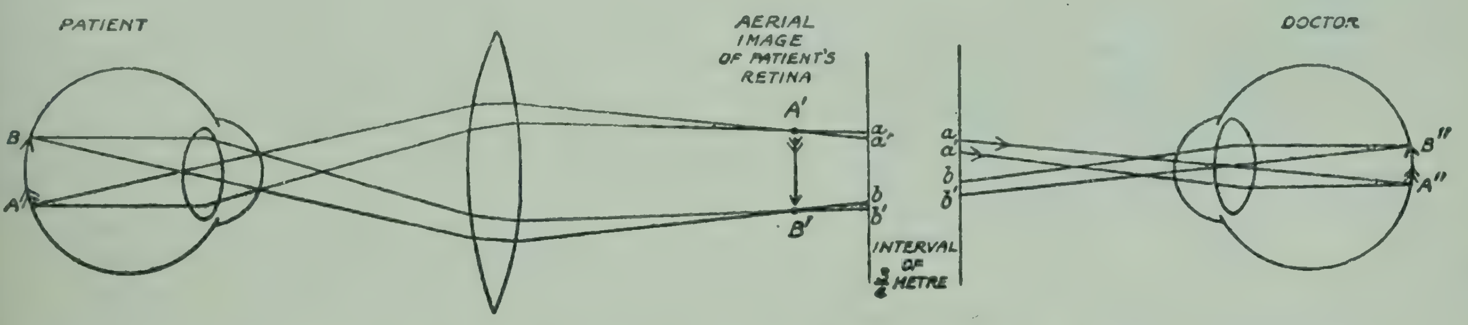


Fig. 278. Diagram to show Paths of Rays from Eye of Patient to Observer when the Indirect Method of Ophthalmoscopy is in use. (Hartridge.)

The image seen by the direct method is erect, that by the indirect is inverted.

On examining the back of the eyeball by either of these methods, the most prominent object is the optic disc or optic nerve papilla, which marks the point of entrance of the optic nerve. It is seen as a pale oval disc surrounded by a deep red background (Fig. 279). From the middle of the papilla the retinal vessels pass into the eyeball, and they are seen diverging from the papilla to ramify over the rest of the retina. The arteries can be distinguished from the veins by their brighter red colour as well as by the stronger reflection of light from their surfaces. The yellow spot is very difficult to see, except in atropinised eyes, since it comes into view only when the observed eye is looking straight into the ophthalmoscope. Under these conditions there is a strong "light reflex," and the pupil contracts up to a pin-point, unless paralysed by means of atropine. In order to see the blind spot, or optic disc, the observed eye must be directed inwards; thus if A is looking at the right eye of B, B must be told to look over A's right shoulder.

By projecting a highly concentrated beam of light on to the side of the eyeball, it is possible to cause sufficient light to pass through the wall for the retina to perceive the stimulus. When that is the case, it is found that the retinal arteries and veins are seen by the subject as dark images on a bright ground (Purkinje's figures). By moving the point of illumination and then measuring the apparent shift of the vessels which occurs, it has been found possible to estimate the depth below the vessels at which the receptive surface of the retina is placed, namely 0·17 to 0·36 mm. Now the average distance between the vessels and the layer of rods and cones is found to be 0·2 to 0·3 mm.; it must therefore be the layer of rods and cones which forms the sensitive layer. The directions taken by the light rays are shown in Fig. 280.

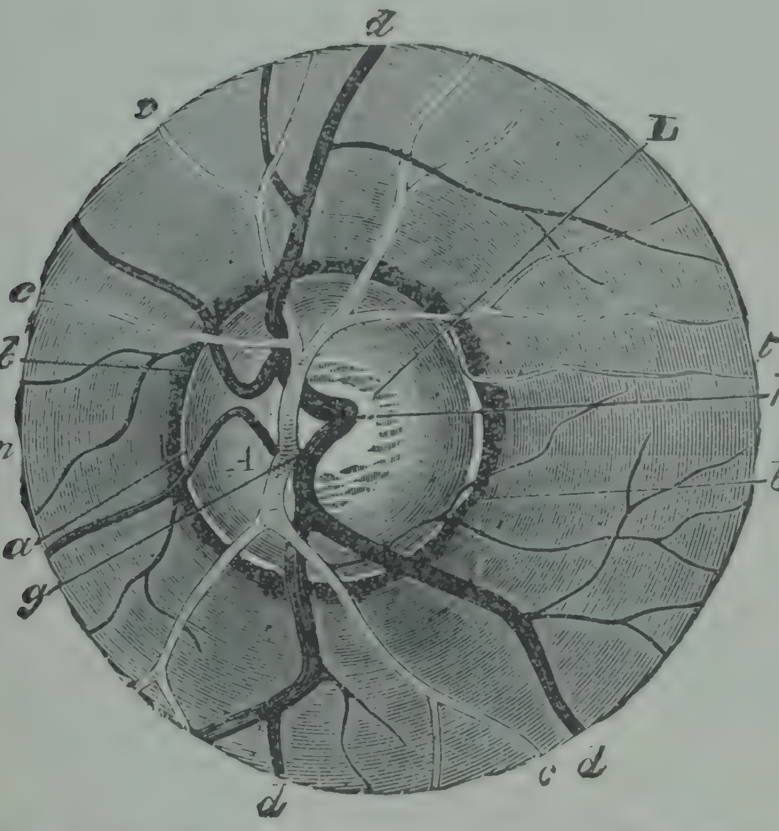


Fig. 279. Ophthalmoscopic View of Fundus of Eye, showing the Optic Disc, or point of entry of the Optic Nerve, with the Retinal Vessels branching from its centre.

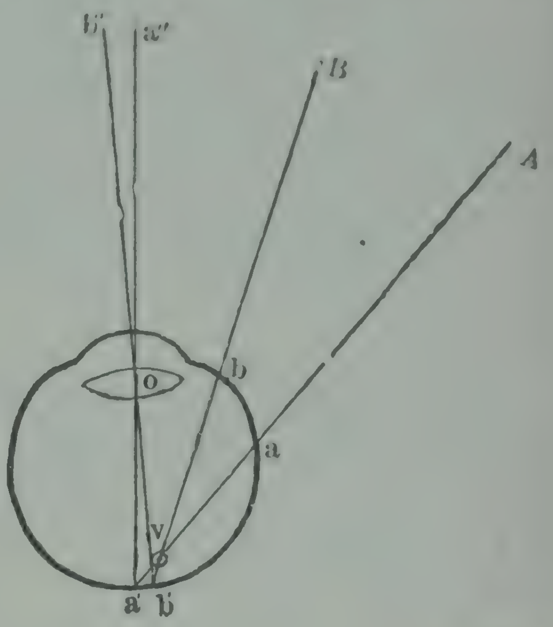


Fig. 280. Diagram of the Path of the Rays of Light in the Formation of Purkinje's Figures.

v represents a retinal vessel. When this is illuminated from A, a shadow is formed on the hinder layers of the retina at a′. This is projected along a line passing through the optic axis, and appears to come from a point (a″) on the wall. On moving the light from A to B, the image of the vessel appears to move from a″ to b″.

Another method of viewing the vessels in one's own eye is to look through a small hole in a metal plate at a smooth white surface. On oscillating the aperture in relationship with the pupil about once a second, the vessels will be seen as shadows on the bright background.

7. THE RELATIONSHIP BETWEEN STIMULUS AND SENSATION

Since we are unable to express our sensations in terms of physical units (we cannot say, for example, when one source is twice as bright as another), two methods of investigation are alone available, namely that which involves the determination of threshold values, and that which depends on the making of comparisons.

In order that a source of light shall be perceived, the image which is formed on the retina must have certain properties. In the first place, it must last for a certain finite time, for if it be of shorter duration than this it will not be perceived. Secondly, it must be larger than a certain size. Thirdly, its intensity must be greater than a certain limiting quantity. Fourthly, the rays it emits must have wavelengths which lie between certain limits. So that in the case of each of these four properties there are limiting values which must be exceeded; these values are called thresholds. The retina has, distributed over its surface, two different types of sensitive organ: the cone apparatus, which has the function of perceiving colours and is used in day vision; and the rod-visual purple apparatus, which is colour blind but very sensitive to light of low intensity and therefore used for twilight vision. Owing to the fact that the distribution of these organs is not uniform, we have to state the part of the retina which is being stimulated, when assigning a value to any of the above-mentioned thresholds. For example, the threshold for intensity may be that which just actuates the rods, that is the achromatic limit, or that which is sufficient to affect the cones and therefore causes an appreciation of colour. Moreover, the value of any one threshold is to a considerable extent controlled by the value of the other factors which have been mentioned ; for example, the time threshold is shorter the greater the size and intensity of the light source. The exact conditions must therefore be carefully stated in quoting the value of a threshold. Lastly, we must consider the personal equation of the observer, and also the state of his vision, for both are affected to an important extent by constitution, health, fatigue, etc.

Intensity Threshold for Light (Achromatic). If a spectrum be gradually reduced in intensity, it loses its colour and finally appears to the eye as a bright band which has its greatest luminosity in the green region of the spectrum and gradually fades towards both the red and violet ends. Since the band is colourless, any one part may be matched by any other part by suitably adjusting the intensities. But compared with the appearance under ordinary intensity, the red region of the spectrum has become greatly reduced in visibility, while the blue has become relatively brighter. The part of the spectrum with maximum luminosity is found to be the yellow when the intensity is high, but to be the green when it is low. It is therefore this shifting of the position of the maximum in the spectrum which has caused red to darken and blue to become lighter. The relative forms and positions of the luminosity (apparent brightness) curves for spectra of various intensities are shown in Fig. 281. If the spectrum of low intensity be still further decreased, a point will be reached at which the different parts become invisible to the eye ; this will occur first

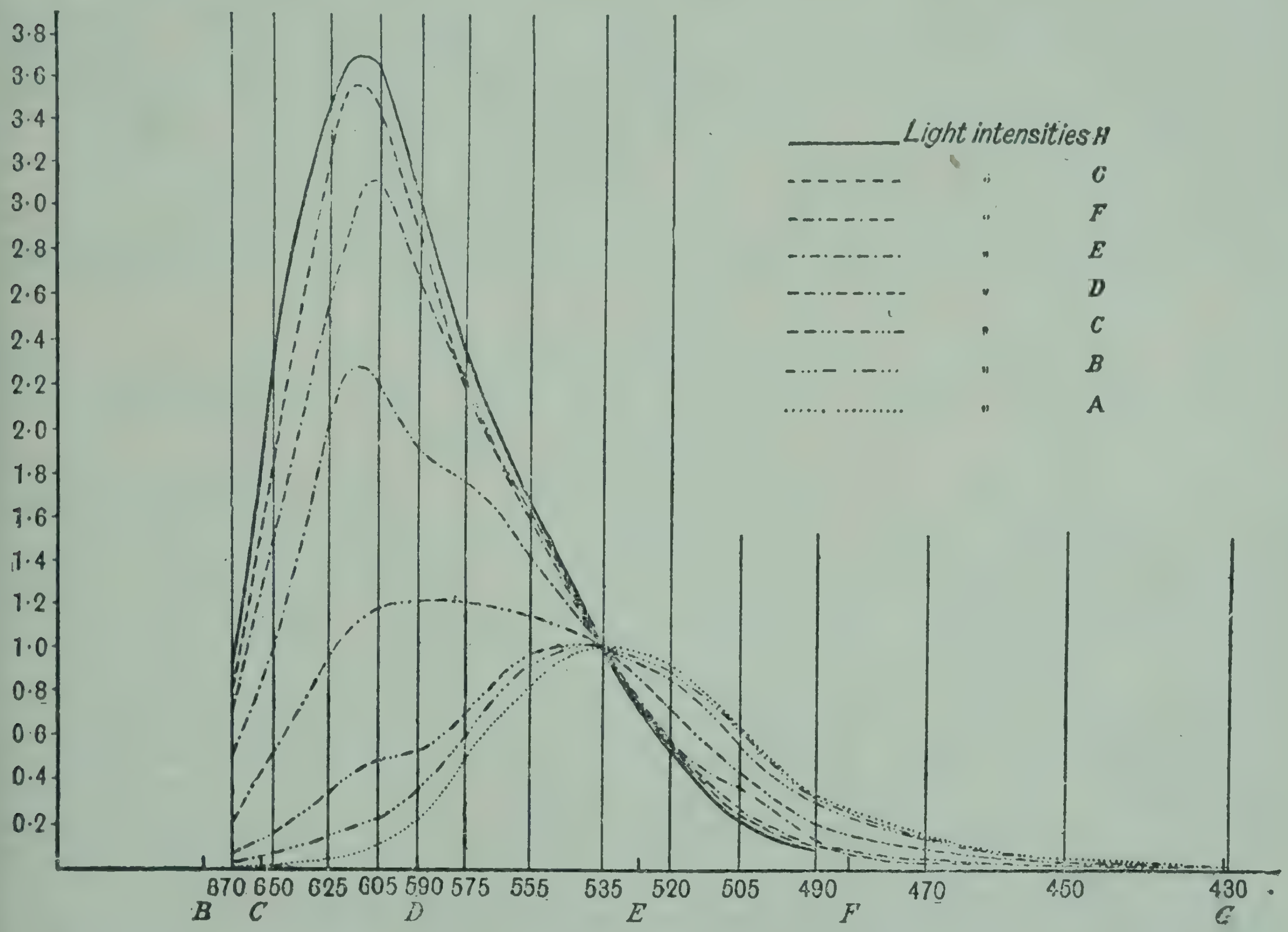


FIG. 281. Luminosity Curves for Spectra of different Intensity. (KÖNIG.)

H=highest ; A=lowest ; Abscissæ=wavelengths ; Ordinates=luminosities. The maximum of curve for light of high intensity is seen to be at 610 $\mu\mu$, that at low intensity 535 $\mu\mu$.

with the ends, and last with the middle (at about 5271 A.U.). It is found, however, that the intensity values at which visibility ceases decrease the longer the eye is kept in the dark—that is to say, the retina gradually becomes " dark adapted." The curves obtained for different degrees of dark adaptation are also shown in Fig. 281.

We must now consider the effects on the achromatic threshold, of size of light source, duration of stimulus and part of retina illuminated. With regard to size, experiment shows that, as the size of the source decreases, so the intensity at which extinction occurs increases, in fact that the area of the source multiplied by its intensity is a constant. With regard to the region of the retina that is stimulated, it is found that the rod-visual purple apparatus is responsible for the appreciation of the light of low intensity ; it is therefore found in all parts of the retina other than the fovea centralis from which rods are absent. The effect of time of stimulus will be considered shortly.

The Directional Sensitivity of the Retina. Stiles and Crawford found that a ray of light reaching the fovea through the centre of the pupil causes a greater stimulating effect than a ray which has reached the fovea through a more peripheral part of the pupil. This effect is not caused by the lens system of the eye and must therefore be due to a property of the retina. Recently Stiles (1939) has repeated the above test using light of different wavelengths and working with both the fovea and the periphery (extra fovea). Some of his results are given in the following table, which gives the light intensities which must be employed in order that a beam of light through the periphery of the pupil shall be as bright to the retina as one passing through the centre of the pupil.

		Nasal Side by			Centre of Pupil.	Temporal Side by		
		3 mm.	2 mm.	1 mm.		1 mm.	2 mm.	3 mm.
Fovea (light or dark adapted)	Red	4·8	2·7	1·5	1·0	0·9	1·1	1·6
	Yellow	3·8	2·3	1·4	1·0	0·9	1·0	1·4
	Green	4·7	2·6	1·4	1·0	0·9	1·1	1·7
	Blue	5·2	3·2	1·5	1·0	0·9	1·1	1·8
Periphery (Para-fovea) (light adapted)	Red	3·8	1·9	1·2	1·0	1·1	1·6	3·0
	Yellow	2·9	1·6	1·1	1·0	1·2	1·7	3·2
	Green	2·8	1·6	1·2	1·0	1·1	1·6	2·6
	Blue	2·8	1·6	1·2	1·0	1·1	1·6	2·6
Periphery (Para-fovea) (dark adapted)	Red	2·9	1·9	1·2	1·0	1·0	1·4	2·3
	Yellow	1·3	1·1	1·0	1·0	1·0	1·1	1·3
	Green	1·0	1·0	1·0	1·0	1·0	1·0	1·0
	Blue	0·9	0·9	1·0	1·0	1·0	1·0	0·9

Thus if two beams of red light fall on the fovea, one through the centre of the pupil, the other through the periphery of the pupil 3 mm. to the nasal side of the centre of the pupil, then for the beams to appear equally bright the peripheral beam must have 4·8 times the intensity of the central one. The conclusion is that light falling obliquely on the retina is less effective than light falling at right angles to the retinal surface.

Intensity Threshold for Colour. If a spectrum of low intensity, which appears colourless to the eye, be gradually increased in brightness, a point will be reached at which the colours begin to be recognisable, first yellow and green, then blue and lastly red and violet. If the intensity at which the colour just vanishes is measured, the curve obtained is similar to that shown in Fig. 282. As the intensity is increased, the point of maximum luminosity gradually shifts from the green to the yellow. As light is gradually increased in intensity, an object is first seen without colour, but on further increase the colour also is recognised; this is called the photo-chromatic interval. It follows from what we have said that the interval is greatest for colours of short wavelength (blue) and least for long (red) (the Purkinje phenomenon). The thresholds for light and colour differ in another important respect: whereas that of light varies with the degree of dark adaptation, that of colour is found by experiment to be nearly constant. With regard to the effect of area of light source, it is found that the same type of relationship exists in the case of colour as for light: that as the area is diminished, so the intensity must be correspondingly increased. The area and intensity, multiplied together, do not, however, equal a constant as they do in the case of light. The appreciation of colour is associated with the cones and is therefore most highly developed at the fovea. As the periphery of the retina is approached, the number of the cones very rapidly decreases, and we should therefore expect to find a limit to the size of the visual field for different colours. This may be tested by means of the perimeter (Fig. 275) and small coloured discs, or more accurately by suitable apparatus for employing spectral colours. By these methods it is found that the colour fields are smaller than those for light, but more or less concentric with them (Fig. 276). The actual size of the fields varies with the intensity and size of the test light source or object. The order in which the colours disappear varies somewhat, but appears to be usually as follows:—First green, then yellow, then red, and lastly blue. The determination of the size of the colour fields is a technique of considerable practical importance, because they are found to become restricted in those progressive lesions of the optic nerves which may finally lead to total blindness, and also in inflammatory conditions of the retina and choroid. Careful examination of the apparent limits of the blind spot by means of similar apparatus was found by Haycraft to show that there is a similar variation in the relative sizes of the fields for different colours (Fig. 283). The same phenomenon is also found round the blind spots formed in the retina by disease.

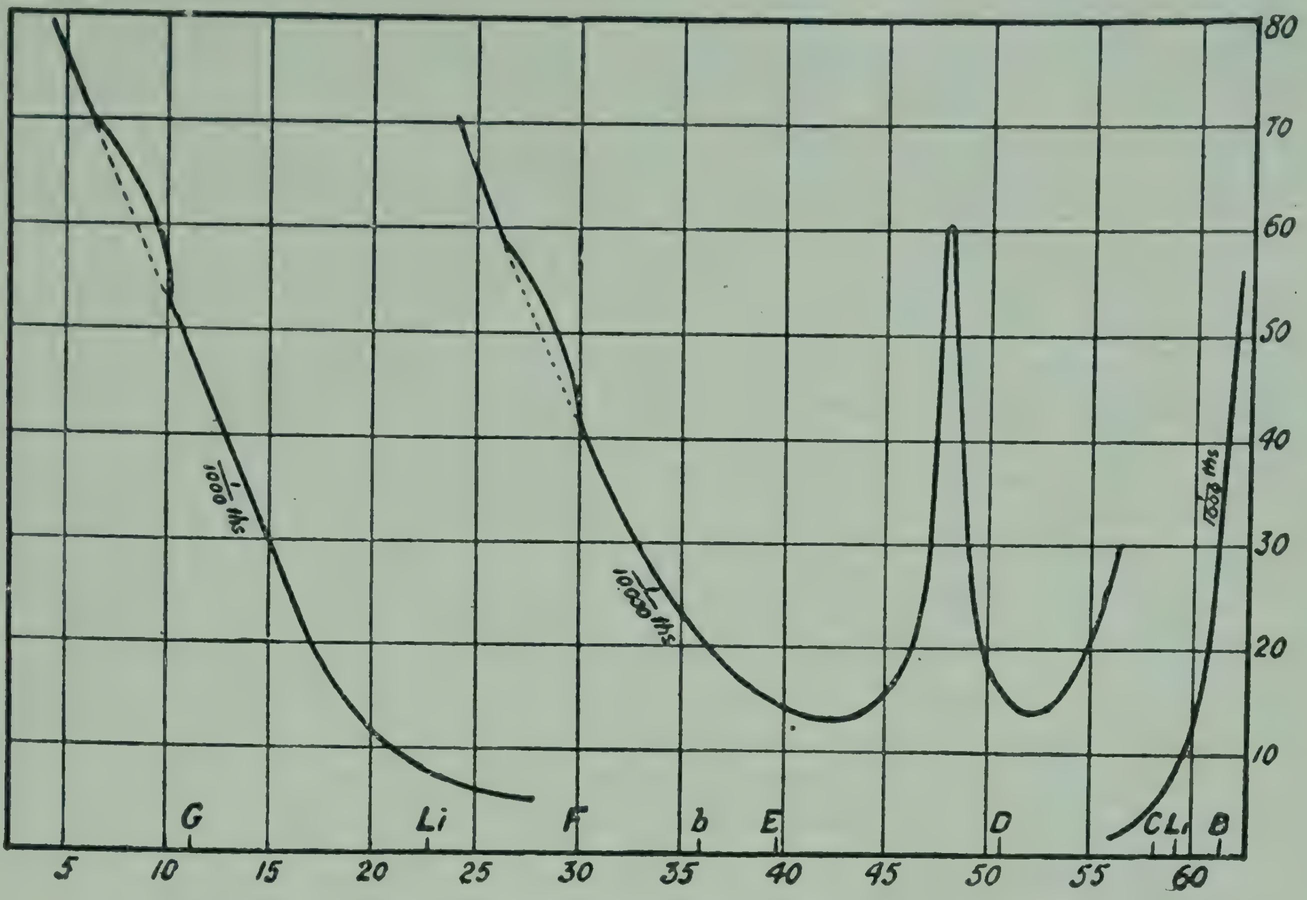


FIG. 282. "Extinction of Colour" Curve. (ABNEY.)
Abscissæ = wavelengths; Ordinates = intensity in candle feet when colour just vanishes.

Size Threshold or Visual Acuity. If a small source of light be gradually reduced in size, a point is soon reached at which it becomes invisible. If it is a coloured source, it as a rule shows a well-marked photochromatic interval, that is, it first loses its colour and then disappears later.

If the intensity of the source is very great, the size has to be much reduced before it becomes invisible: it is because of this that we see the stars. If the size and intensity at the point of disappearance be measured, it is found that when multiplied together they give a constant, so that as in the case of the light threshold, the determining factor appears to be the amount of light which falls on the retina. In the case of intermittent illumination a similar relationship is found.

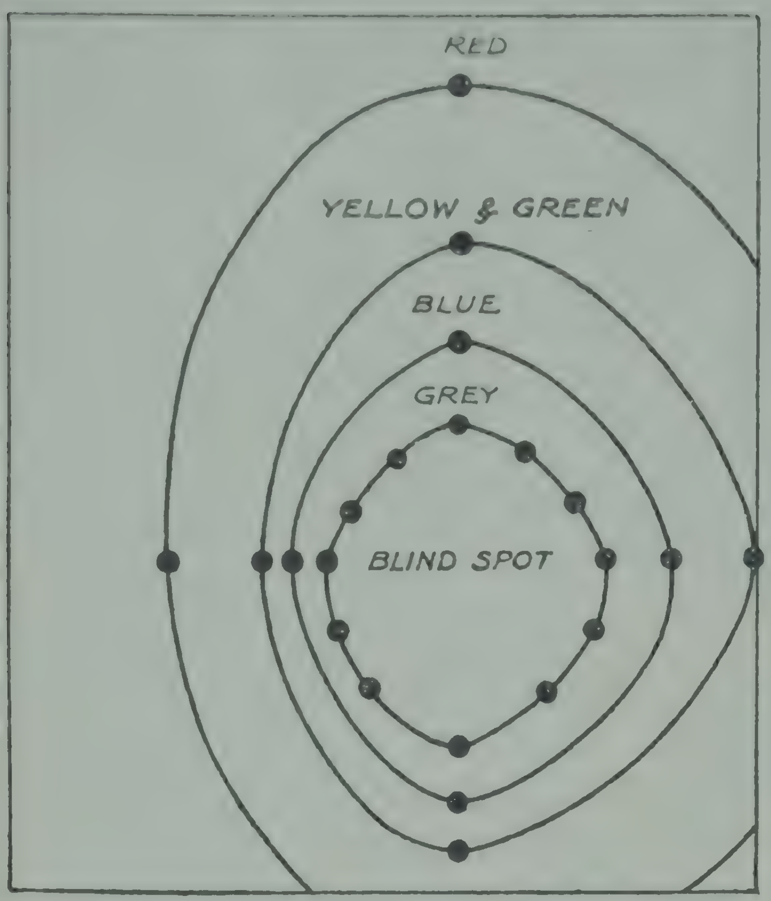


FIG. 283. Limitation of Colour Fields round the Blind Spot. (HAYCRAFT.)

Visual acuity is the ability to see as separate the images of small bright light sources of any shape placed very close together. Experiment shows that the distance between the sources must be increased as their distance from the eye is increased: in other words, that the angle which they make at the eye must be greater than a certain limiting value. The angle usually obtained is an angle of one minute, and on this the test types used in practice for testing the visual acuity of patients is based (see p. 396). Persons with exceptionally good vision are able to see the separate images, when the angle which the sources make at the eye is very much less than this, namely 24 seconds. Assuming that the posterior nodal point is 15·6 mm. from the retina (this being the distance in the normal eye), 24 seconds corresponds to a distance between the images at the retina of 0·0018 mm. The diameter of the cones is between 0·0020 and 0·0030 mm., and in the fovea they are very closely packed. The maximum visual acuity is therefore certainly as great as the size of the cones would lead us to suppose possible.

A dark spot on a bright ground is similar to the case just considered, because for the dark spot to be recognised it must subtend at the eye the minimal angle mentioned above. But increasing the intensity of the ground or the blackness of the spot will only make a very small difference, so that unlike the bright spot, an increase in the intensity does not make up for a difference of size.

The Dependence of Visual Acuity on Light Intensity. It is well known that visual acuity is very low when the eye is used under conditions of dim illumination. Even well-defined objects appear nebulous when seen in poor light. As the illumination is increased, definition improves with very great rapidity until at great intensities of illumination very high acuity is reached. When visual acuity is plotted against illumination an S-shaped curve is usually obtained.* Some observers, *e.g.* König, have reached the conclusion that visual acuity reaches a maximum value at about 100 ft. candles and that no improvement results from increasing illumination above this amount. Lythgoe, however, states that if care be taken to adapt the eye adequately, then visual acuity goes on improving as the illumination increases above 100 ft. candles. This relationship of visual acuity and intensity forms the basis of modern methods of artificial illumination, the brightness used depending on the visual acuity that is required of the observer. Thus, for passages and store-rooms 1/10 ft. candle may be sufficient. For coarse work not requiring detailed examination, 1 ft. candle may suffice. For clerical work, 10 ft. candles may b found necessary. For work where fine detail is to be seen, 100 ft. candles may be wanted, and 1,000 ft. candles will probably not be excessive for the finest work of all. Thus, for seeing very fine structures under a microscope, a bright microscope field is necessary.

---

* "Foundations of Experimental Psychology," 1929, p. 246, Fig. 15.

In order to double the visual acuity it is usually necessary to increase tenfold the intensity of illumination of the objects looked at, and their surroundings. The beneficial results of using bright lights appear to be due to two effects: (1) An increase in the discriminating power of the retina; (2) a decrease in the size of the pupil. The latter improves the definition of the image on the retina produced by the lens system of the eye, and at the same time increases the depth of focus, that is, it enables the eye to observe sharply objects at more than one distance from the eye. Two examples may make these points clear: A man mending a watch, and a woman sewing. When a watch is opened it will be seen that the works consist of two metal plates between which the gear wheels are pivoted. These are at different distances from the observer, in consequence if he focusses a gear wheel sharply the plates above or below it tend to be out of focus. The smaller the pupil, the greater the sharpness of definition of these plates owing to the increased depth of focus. Further, each bright metal surface reflects the light strongly, and therefore tends to look larger than its true size, because of the irradiation which is caused by the refractive errors of the peripheral parts of the lens system of the eye. The smaller the pupil the less the irradiation, and the more nearly do objects appear the correct size. An increase in the illumination causes the pupil to contract, and thus brings about an improvement both in the depth of focus and in irradiation. Sometimes in watchmaking and repairing it is necessary to detect extremely fine objects, such as specks of dirt or irregularities in the teeth of gear wheels. In this case the use of high illumination by reducing irradiation, and improving the defining power of the retina greatly increases the powers of observation. In the case of sewing, the material, unless it is stretched quite flat by the use of an embroidery frame, lies at different distances from the observer. The needle and thread is normally closer to the observer than any part of the material that is being sewn. In order to see needle, thread and material sharply at the same time, considerable depth of focus is required and this necessitates the use of a small pupil such as is obtained by illuminating the material brightly. Further, when the thread is extremely fine and great precision is required, for example when fashioning silk stockings by hand, then high visual acuity is necessary with the minimum amount of irradiation. This difficult work is found to be greatly facilitated by the use of high intensities of illumination. When the objects to be looked at are themselves dark in colour, for instance oxidised metallic surfaces or dark fabrics, then even greater light intensities should be made use of. Direct sunlight is not too bright for such purposes.

The Dependence of the Brightness Difference Threshold on Light Intensity. The ability of the eye to detect a difference of brightness is found to change considerably with illumination. When illumination is poor, quite big differences of brightness are not noticed. As illumination improves the differences which can be perceived get smaller and smaller until at very high illuminations, the differences become extremely small. If the difference, which can be detected by the eye, is plotted against the intensity, a curve* is obtained from which the conclusion is that, as in the case of visual acuity, the illumination required by an observer will depend on the differences of brightness that have to be visible to his eyes. If the differences of brightness are considerable, for example white objects lying on a black ground, a low degree of illumination will suffice. As the differences get smaller, the illumination required gets larger, so that where it is necessary to see objects which have nearly the same brightness as their surroundings then quite high intensities of illumination may be required.

---

* "Foundations of Experimental Psychology," 1929, p. 254, Fig. 18.

Much the same statement applies to differences of colour. The perception of small colour differences is very greatly facilitated by a high degree of illumination. In consequence, for the most perfect colour matching which the eye can carry out, a high degree of illumination should be employed.

Hecht's Theory of Vision. Hecht's theory is based on the duplicity theory, namely that the retinal receptors consist of both rods and cones; the rods occupying the more peripheral parts of the retina and being excluded from the fovea; the cones occupying the central part of the retina and having exclusive tenancy in the fovea. According to Hecht, these form two populations which differ considerably in the quantity of light required to stimulate them, the rods being, on the average, enormously more sensitive than the cones. Now according to Hecht the rod population consists of members which differ greatly in sensitiveness; a few have relatively low sensitivity, a few have relatively high sensitivity, and between the two come the main mass of the rod population. A similar statement, according to Hecht, applies to the cones. Now these two populations overlap so that as illumination is increased from zero, and therefore more and more rods are brought into play, some cones have commenced to act before the intensity has reached such a value as to cause a response from the least sensitive of the rods. This hypothesis is advanced by Hecht in order to explain the effects of increasing illumination on visual acuity and on brightness difference threshold. Thus, at very low intensities, only a few scattered rods respond, the consequence being low visual acuity and low discrimination of difference of brightness. A small increase in illumination will bring more rods into an effective condition, visual acuity will thus improve and, since the additional rods thus brought into play do not differ from one another so much in sensitiveness as the ones stimulated by the previous low degree of illumination, the discrimination of differences of brightness will also show an improvement. As the illumination is still further increased, more and more rods act until the brightness reaches a value at which the main rod population begins to function. At this point the difference of threshold between these individual rods is small, and so the ability to perceive differences of brightness has reached a much higher value than before. Now as the illumination is still further increased the cone population begins to be affected. The number of percipient elements, therefore, increase with a corresponding increase in visual acuity and, since the cones are more nearly alike in sensitiveness than the rods, the ability to discriminate small differences of brightness is also greatly improved. Further increases in illumination bring more and more cones into play until the stage is reached at which the main cone population begins to come in, and this causes an enormous increase in visual acuity and a correspondingly large improvement in the ability of the eye to perceive small differences of brightness.

There are three criticisms which can be advanced against this theory: (1) that it requires very high illumination indeed to cause the whole of the cone population to reach a functioning state. Which means that under ordinary conditions of illumination only a fraction of the cones are playing a part in vision. It may ultimately be found that such is actually the true state of affairs just as there may be sensory organs for pain which, for the greater part of our lives, are never called on to respond, but which are capable of functioning immediately a nocuous stimulus of adequate intensity is applied to the particular part of the body in which these pain end organs are situated. (2) That it assumes that the sensory structures of the retina obey the all or none law; that is that a rod or a cone can only be in one of two conditions, not responding at all or responding to the greatest extent of which it is capable. If the rods and cones obeyed this law they would differ from all the other sensory end organs which have been experimentally investigated. Moreover, experiments on the retina of Limulus have shown that its retinal receptors are graded in their responses. When illumination is low only a few impulses per second are transmitted to the brain by the optic nerve. As illumination increases, the number of impulses sent up the optic nerve increases.* In human beings, therefore, we should expect the impulses transmitted by the various fibres of the optic nerve to depend on the intensity of stimulation of the rods and cones to which the various fibres are attached. (3) That it does not consider cases where the visual acuity exceeds by a very material amount that which would be expected according to the law of geometrical optics. The position with regard to these cases will now be examined.

The Diameter of the Cones of the Fovea. A large number of investigators have determined the dimensions of the cones taken from the fovea centralis of the eye. They vary considerably as the following table shows. The mean value being about $3{\cdot}2\mu$.

Observer.	Diameter of cones.
Max Schultze	$2{\cdot}8\mu$
H. Muller	less than $3{\cdot}0\mu$
Merkel	$3{\cdot}0\mu$
Welcker	$3{\cdot}1$–$3{\cdot}5\mu$
Wadsworth	$2{\cdot}5\mu$
Kühnt	$2{\cdot}0$–$2{\cdot}5\mu$
Kölliker	$4{\cdot}5$–$5{\cdot}4\mu$
Koster	$4{\cdot}4\mu$
Greeff	$2{\cdot}5\mu$
Dimmer	$3{\cdot}0$–$3{\cdot}5\mu$
Mean	$3{\cdot}2\mu$

If the focal length of the eye be assumed to be 15 mm., then $3{\cdot}2\mu$ corresponds to an angle of 44 seconds of arc.

The Resolving Power of the Human Eye. The values obtained for the resolving power of the eye differ according to the type of test which is used. Thus, Hooke found that two stars must be separated by approximately 60 seconds of arc in order that they may be seen separately by the eye. The distance between the bright bars of a grating, in order that resolution shall occur, have been found in general somewhat less, as shown in the following table.

Observer.	Value at retina in		Angle in seconds of arc.
Lister	$4{\cdot}6\ \mu$	...	64
Hirschmann	$3{\cdot}6\ \mu$	...	50
Bergmann	$3{\cdot}8\ \mu$	...	52
Helmholtz	$4{\cdot}6\ \mu$	...	64
Uthoff	$4{\cdot}0\ \mu$	...	56
Cobb	$4{\cdot}6\ \mu$	...	64
Mean	$4{\cdot}2\ \mu$	...	58

For some other test objects the acuity of the eye exceeds, by a considerable margin, the above values. Thus, Volkmann, in 1863, found that the smallest difference that can be perceived between the breadth of two white objects on a black ground was as little as 10 seconds of arc. Wulfling, in 1892, confirmed Volkmann's observations. Hering, in 1899, showed that the edges of two rectangles did not appear in line with one another if they were shifted sideways by as little as 10 seconds of arc. He named this ability to identify small displacements as "the position sense." Hartridge confirmed Hering's value. Bryan and Baker, 1912, found that one line of a vernier can be set into coincidence with another line of a vernier with an accuracy of between 8 and 10 econds of arc. Hartridge found that objects with diffuse contours (like bsorption bands) can be set into line with an accuracy of approximately 1 seconds of arc. It will be observed that all these experiments indicate an ccuracy of observation of approximately one-sixth of the diameter of a foveal one. Even greater accuracy is found when the acuity of vision is determined for the narrowest black line that can be perceived on a white ground. The values obtained for several observers is given in the following table.

---

* *J. Cellular & Compar. Physiol.* (1), 1932, p. 277, Fig. 6 (p. 291).

	Seconds of arc.
Aubert	6
Smith Kastner	3·5 Spider's web.
Hartridge	3·6 Bright brass wire.*
Hartridge and Owen	3.1 Black brass wire.*

Recently Hecht and Mintz, using as test object a black line whose length was many times its width, obtained values for visual acuity even better than those quoted in the above table. They obtained values as low as 0·5 seconds of arc.

The acuity of stereoscopic vision has also been found to be extremely high. Thus, Bowden found an accuracy of 5 seconds of arc, while Breton, 1920, found an even greater accuracy of 4 seconds of arc ; which would correspond to an accuracy of one-fifteenth of the diameter of a foveal cone. It is clear that some theory is required which will account for the very great acuity that experiment shows that the eye possesses for these kinds of test.

Hartridge's Theory of Visual Acuity. On the basis that the eye suffers from chromatic aberration and diffraction, Hartridge calculated the distribution of light intensity in the image formed on the human fovea. This calculation was done for (*a*) a grating test object, (*b*) for a narrow black line on a white ground, (*c*) for the coincidence of two black lines on a white ground (or two white lines on a black ground), and (*d*) for the stereoscopic sense of position. From the distribution of light thus calculated, it was possible to ascertain the average illumination falling on the cones of the retina. It was found, for example, in the case of a single black line on a white ground, that that cone which according to geometrical optics corresponds with the centre of the black line, receives 13 per cent. less light than the cones on either side of it. If, then, this cone sends impulses up the optic nerve fibre connected to it which are noticeably less in number than the impulses on the cones on either side, the existence of this narrow black line will be identifiable and its position in space can then be observed. Similar calculations for the other types of test mentioned above have given the values shown in the following table.

A. Cases in which there is a constant difference of intensity without eye movements.	
Single black line on white ground	13 per cent.
Grating	20 ,,
Double line	12 ,,
B. Cases in which there is a variation of intensity when movements of the eyes occur.	
Coincidence of two lines	22 per cent.
Position of edge of contour	32 ,,
Stereoscopic vision	12 ,,

If the figures on the table above be examined, it will be noticed that the values obtained when there is a constant difference to be observed are somewhat smaller than when with movements of the eyes a comparison has to be made between the intensity at one moment and that at another. Otherwise these values are fairly concordant if allowance is made for experimental error, for the fact that the observations of the different test objects have been made by different observers, and for the difficulty of making even approximate calculations of the intensity of light falling on the retina.

We may summarise what has been stated above as follows : for certain kinds of test object the acuity of the eye is such as to correspond to the employment of an area considerably less than that represented by the diameter of a cone. This acuity can be accounted for if the intensity of light falling on a certain cone is noticeably different to that on the cones on either side of it. The illumination on the central cone must be from 12 per cent. to 32 per cent. greater or less than that on the two cones next to the central cone.

---

* Both mounted in front of white paper at such a distance that no shadow was cast.

Now the question which presents itself is this: "is there any evidence that such differences of intensity can be observed?" because, if they can, then the above theory is substantiated.

Now, when large retinal areas are being stimulated and the intensity of illumination is high, the smallest difference of intensity that can be perceived as given by various observers is as follows:—

König	1·56 per cent.
Helmholtz	0·6 ,,
Aubert	0·68 ,,
Hartridge (with absorption bands)	about 0·5 ,,

But experiment shows that the smaller the area of retina that is being tested, the greater must be a difference in intensity for it to be perceived. For so small a retinal area as that corresponding to one or two cones it would be quite possible that the difference threshold would have fallen from the high values given in the table above to the lower values of 12 per cent. to 32 per cent. required to explain the acuity of the eye mentioned previously in this section.

Hecht and Mintz, employing a similar method of calculation, found for a black line test object a difference of illumination of 0·95 per cent. between the row of cones corresponding to the centre of the geometrical image of the black line and the rows of cones on either side of the central row.

Now all the above experiments were probably performed under approximately ideal conditions of illumination. As the intensity of illumination decreases there will be a deterioration in the performance of the eye which is due to the fact that the ability of the eye to detect a small difference of brightness depends on adequate illumination, and falls progressively as illumination decreases.

Time Threshold. In considering the time threshold, two different sets of conditions require description: firstly, the minimum time during which a given stimulus must act in order to reach consciousness, and secondly, the minimum rate at which a series of stimuli must follow one another in order to give a uniform impression without flicker. Both are of considerable importance, since the first enters into such problems as the determination of the length of time during which a lighthouse beam should be projected in a given direction; the second because it gives a reliable method of comparing the intensity of lights of different colour. Experimental investigation of the first type of time threshold is effected by measuring the length of stimulus necessary to cause a source of a certain intensity to affect the retina; and it is found that the lower the intensity the longer must the image fall on the retina. But if the eye be dark adapted, if the time and the intensity values be multiplied together, then within limits a constant is obtained. On the other hand, in the light adapted eye, the value is found to vary somewhat, but is sufficiently constant to show that the relation between intensity and time is approximately the same. Within limits, therefore, we find that at the threshold the total amount of light is constant, whether it be of high intensity for a very short period of time, or of low intensity for a correspondingly longer one. This relationship ceases to be true if the time of stimulation is longer than about one-tenth of a second, and this is apparently due to the fact that the retinal apparatus reaches a steady state in about one-quarter of a second in the dark adapted eye (rod-visual purple apparatus). A lighthouse flash should therefore be visible to the eye for this length of time, in order to make the greatest possible impression. For coloured lights, approximately the same values are found, provided that allowance is made for the comparative intensity of the colour. Since the time required for the retina to reach a steady state is nearly that at which a series of stimuli must fall on the retina in order to produce a uniform sensation, for intensities near the threshold the rate at which flicker disappears is one stimulus every quarter of a second. But it is found that as the intensity rises, the rate must be increased in order to abolish flicker. The rule which most nearly expresses the relation appears to be that geometrical increase in the intensity requires an arithmetical increase in the rate. Sherrington showed, however, that the results are affected by simultaneous contrast. This phenomenon of flicker, as has been said, is used in practice for measuring the intensities of light sources. Two methods are employed: firstly, that in which the two light sources to be compared are measured separately for the intensity at which flicker ceases when the same rate of stimulation is used for both; and secondly, that in which the two sources are caused to fall alternately on the eye, and are adjusted in intensity until flicker ceases. Of the two methods, the latter is the more accurate. The value of these methods lies in the fact that they measure brightness independently of colour. The shape of the curves obtained by plotting the luminosity of different parts of the spectrum has been shown in Fig. 281.

Lastly, we have to consider the relationship between time, intensity and apparent brightness, in the case of an intermittent source, the rate of which is sufficiently great to avoid flicker. Experiment shows that the brightness increases in proportion as either the intensity or the time intervals are increased, and further that equal brightness is obtained if the time multiplied by the intensity is constant. This statement is true of both the cone and the rod apparatus, and is known as the Talbot Plateau Law. Use is made of this law in the sector method of controlling intensity (see p. 424) because the intensity is proportional to the time during which the light is allowed to pass through, which is in its turn controlled by the angle between the blades of the sector.

Colour Threshold. On testing the violet end with a photographic process, or the red end by a thermopile, it can readily be shown that the spectrum extends at both ends far beyond the visible limit. The visible limit at the red end under the most favourable conditions has been found to be 8350 A.U., while under ordinary circumstances it is difficult to go beyond 8000 A.U. Since rays beyond this reach the retina in considerable amount, the limit cannot be caused by opacity of the eye media, and must therefore be due to an actual inability on the part of the retina to record their presence.

Of several hypotheses which might be advanced for this inability, the most probable is that the retinal pigments are unable to absorb rays in the infra-red part of the spectrum, and therefore, according to Draper's law, such rays cannot produce photochemical change and cannot be perceived by the eye. It may be mentioned in this connection that few organic pigments absorb strongly in the infra-red.

The limit at the violet end is less easy to determine, because the eye media, in common with a large number of other bodies, have the property of fluorescing when the ultra-violet rays fall on them, *i.e.* they convert them into rays of longer wavelength and therefore make them visible. The resulting impression is, however, quite different, because, since these rays are generated in the eye media themselves, they are spread over the retina as a haze without there being any proper image formation. The limit of the visibility of the violet end of the spectrum appears to be at about 3,800 A.U., while the portion which is seen because of the fluorescence, which it produces, and which appears a pale lavender, ends at about 3,200 A.U. Since the wavelength of the extreme red rays is a little more than double that of the extreme violet, the eye is sensitive to a little over an octave. The range of appreciation of the eye is therefore very much smaller than that of the ear, which is about 10 octaves. As age increases, the eye media become yellow in colour; this change particularly affecting the lens, the violet end of the spectrum becomes shortened owing to absorption. On removing the lens of the eye, as in an operation for cataract the sensitiveness to the violet end of the spectrum is considerably increased It would therefore seem certain that the limitation of the spectrum at the viole end is largely due to absorption by the eye media, and not to inappreciatio on the part of the retina. The causes of the limitation of the two ends of the spectrum are therefore different.

Difference Thresholds. Beside the thresholds for light, colour, time and wavelength, which may be called absolute thresholds, there are certain difference thresholds that must be considered. Thus, for example, a certain finite difference must exist between the intensities of two sources of light of the same colour for a difference between them to be appreciated by the eye. There are four principal types of difference threshold: that of intensity, that of colour, that of saturation, and that of size.

Difference Threshold of Intensity. It is found by experiment that a just perceptible difference between the intensities of two surfaces varies with the mean value of their intensities. Thus, supposing that it had been found by one experiment that a difference of intensity of one foot candle was necessary in order that two sources should be just distinguishable, the average intensities of which were one hundred foot candles, then in another case in which the average was 25 F.C. the least perceptible difference would be found to be one-quarter F.C. This is one illustration of the principle known as Weber's law. It appears to be true for light of medium intensity, and for sources not separated by more than a small interval. But the least perceptible difference is found by most observers to be less than that taken for purposes of illustration above, namely one-hundreth part of the mean intensity; thus Helmholtz found it to be a $\frac{1}{167}$th, other observers have obtained even smaller fractions. It is interesting to find that the results are not influenced by the size of the pupil.

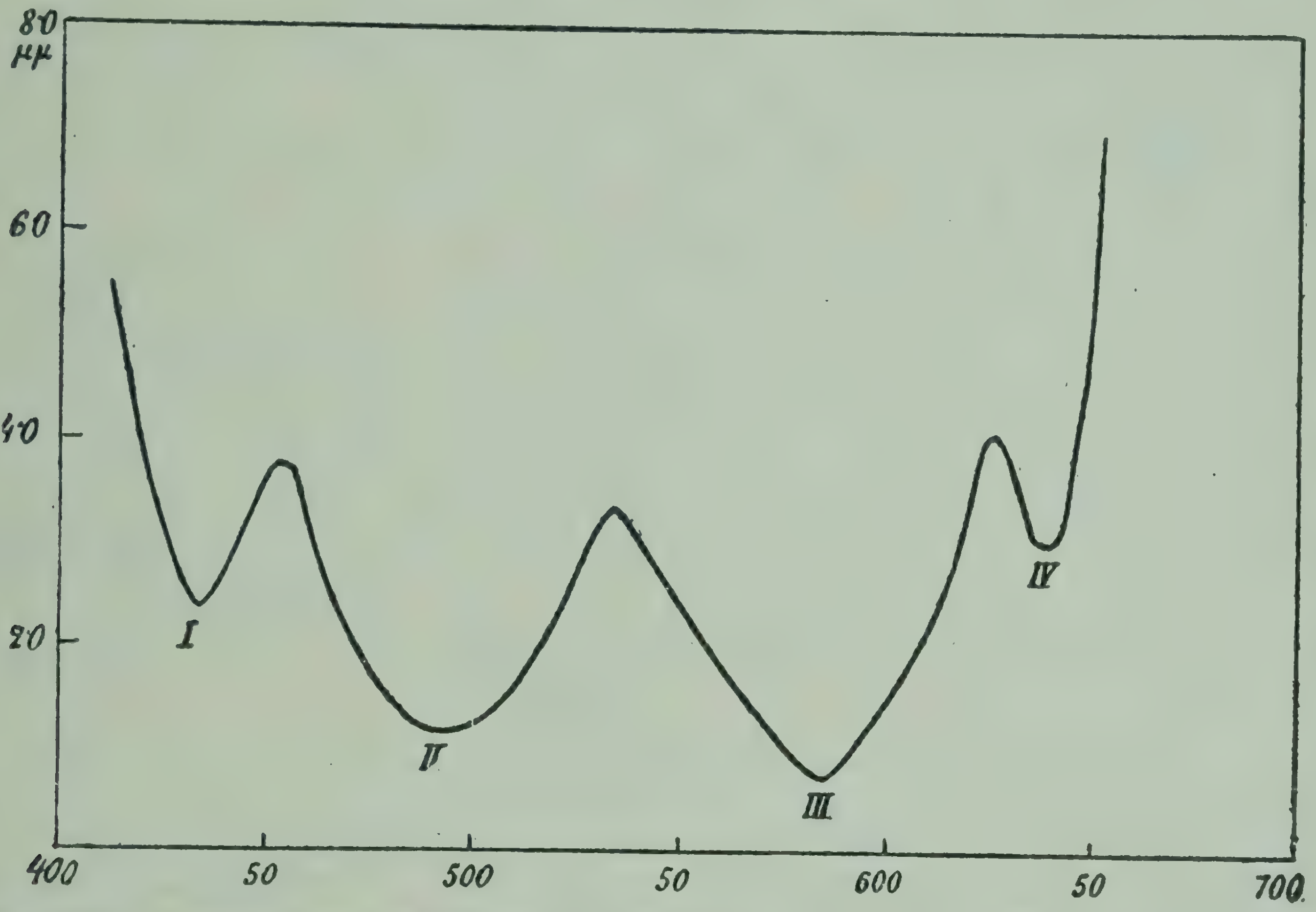


Fig. 284. Curve showing Difference Threshold of Colour at different parts of the Spectrum.

Abscissæ = wavelengths. Ordinates = difference between wavelengths capable of being discriminated. (Steindler.)

Difference Threshold of Colour. If the range of colours exhibited in the spectrum be carefully examined, it will be seen that there are certain parts, notably at the red and violet ends, at which the change of colour with wavelength is a very gradual one. At other parts, on the contrary, the change of hue is very rapid; the yellow region at 5,800 A.U. may be given as example. If therefore we determine by experiment what difference of wavelength is just perceivable by the eye, we find that it varies with the part of the spectrum under observation. We may therefore conveniently express the difference threshold in different parts of the spectrum in the form of a curve, as in Fig. 284. In persons with normal vision the total number of different hues in the spectrum is calculated to be 165. In persons with colour blindness the number is greatly reduced.

This fact has been applied by Edridge Green for the detection of colour blindness. It should be pointed out that, since in this method the spectrum itself is presented to the observer, so that there is a gradual change from one colour to the next, it is the threshold of rate of change of colour that is determined, and not difference threshold of colour. Edridge Green's method gives only from 21 to 28 monochromatic patches in the spectrum.

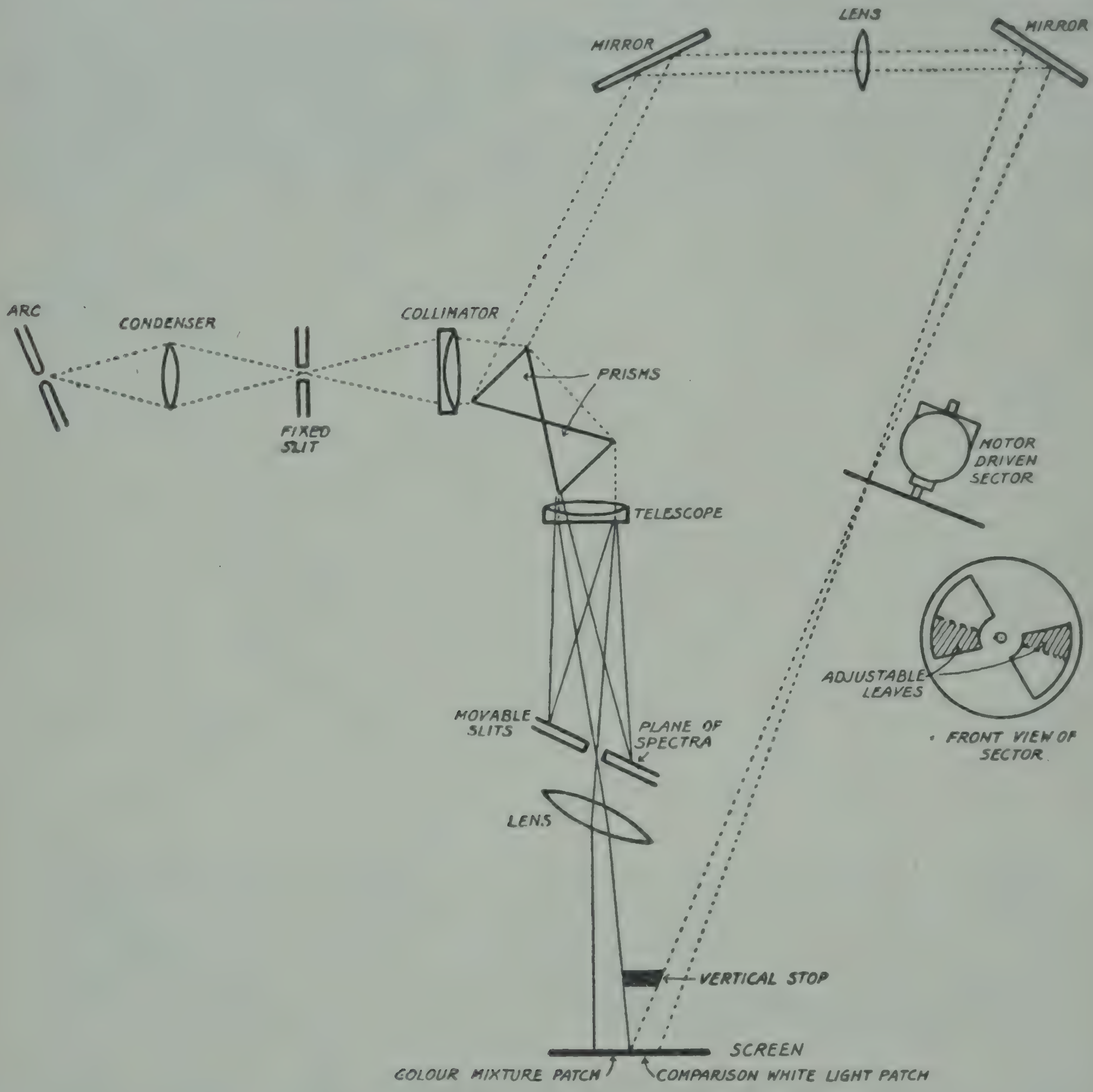

LENS
MIRROR
MIRROR
ARC
CONDENSER
COLLIMATOR
PRISMS
FIXED SLIT
MOTOR DRIVEN SECTOR
TELESCOPE
ADJUSTABLE LEAVES
• FRONT VIEW OF SECTOR
MOVABLE SLITS
PLANE OF SPECTRA
LENS
VERTICAL STOP
SCREEN
COLOUR MIXTURE PATCH
COMPARISON WHITE LIGHT PATCH

Fig. 285. Colour Patch Apparatus or Mixture and Comparison of pure Spectral Colours. (Abney.)

Difference Threshold of Saturation. By saturation is meant the amount of white light which is present with, and is therefore diluting, a colour. The threshold would appear to be of the same order as that of intensity given above, namely that a difference in the amount of white light diluting a colour by $\frac{1}{180}$th of the total intensity present can be just appreciated by the eye.

Difference Threshold of Size. If two objects are the same distance from the eye, and are close to one another and in similar positions, a difference of one-hundredth the mean size can, as a rule, be appreciated. If they are at different distances from the eye, or are far apart, or are not in similar position (*e.g.* one perpendicular and the other horizontal), then considerable errors may occur.

The Flicker Method. The majority of observers can obtain consistent measurements of intensity with the above method only when the colours of the two patches are exactly alike. Thus it is difficult to adjust a green light to be of equal apparent brightness (luminosity) with a red because the difference in colour makes the judgment of brightness inaccurate. Abney found that in his own case practice greatly increased the certainty with which the measurement could be made. A more reliable technique is given by the flicker method. The two patches are viewed through a rotating sector, the speed of rotation of which can be controlled. The intensity of one of them is now adjusted so that, when the speed of the sector is altered, both commence to show and to cease showing flicker at the same time. By applying this method to the colour mixing apparatus, the luminosity of the different parts of the spectrum can be determined. The curves obtained at different luminosities have already been given in Fig. 281. Tested by this method, different observers show individual peculiarities, which amount in some cases to a greatly diminished perception of a certain part of the spectrum. Some of the types met with will be described later.

The Method of Comparisons. In the application of this method three different series of investigations have been carried out: (1) To determine the intensities of three primary colours which, when mixed together, will match the different spectral colours or white light. (2) To determine the intensities and wavelengths of the complementary colours. (3) To ascertain the intensities and wavelengths of red and green rays which, when mixed together, match a pure spectral yellow.

A colorimeter of some form or other is used for these tests. Abney's apparatus may be described as a typical example. Light from an arc lamp is focussed on to the slit of a powerful spectroscope, which consists of a collimator, a train of prisms and a telescope. The spectrum thus produced is caused to fall on three slits, one of which corresponds with the red, another with the green, and the third with the blue. The light having passed through the slits falls on a lens which forms an image of the prism faces on a screen. The light from the slits thus recombines on the screen to produce a bright patch, the colour of which alters according to the intensities of the three components. To one side of the patch a second patch of light can be thrown from the arc lamp, and this also may be varied in different ways according to the nature of the experiment. The intensities of the different beams could be modified by altering the widths of the slits; a preferable method is to employ rotating sectors, the angles between the blades of which can be varied at will (see Fig. 285).

Colour mixing experiments performed with this apparatus give results that have already been briefly considered in Section I. It is found that not only do the three primary colours, when mixed together in the right proportions, form a white light that is indistinguishable from ordinary white light, but they can also be made to match the whole range of colours both of the spectrum and of pigments. It was also described how that certain pairs of colours when mixed in the right proportions are able to match white, and that these pairs are called complementary colours. If the colours that are mixed are farther apart than are the complementary colours, then the mixed colour is found to

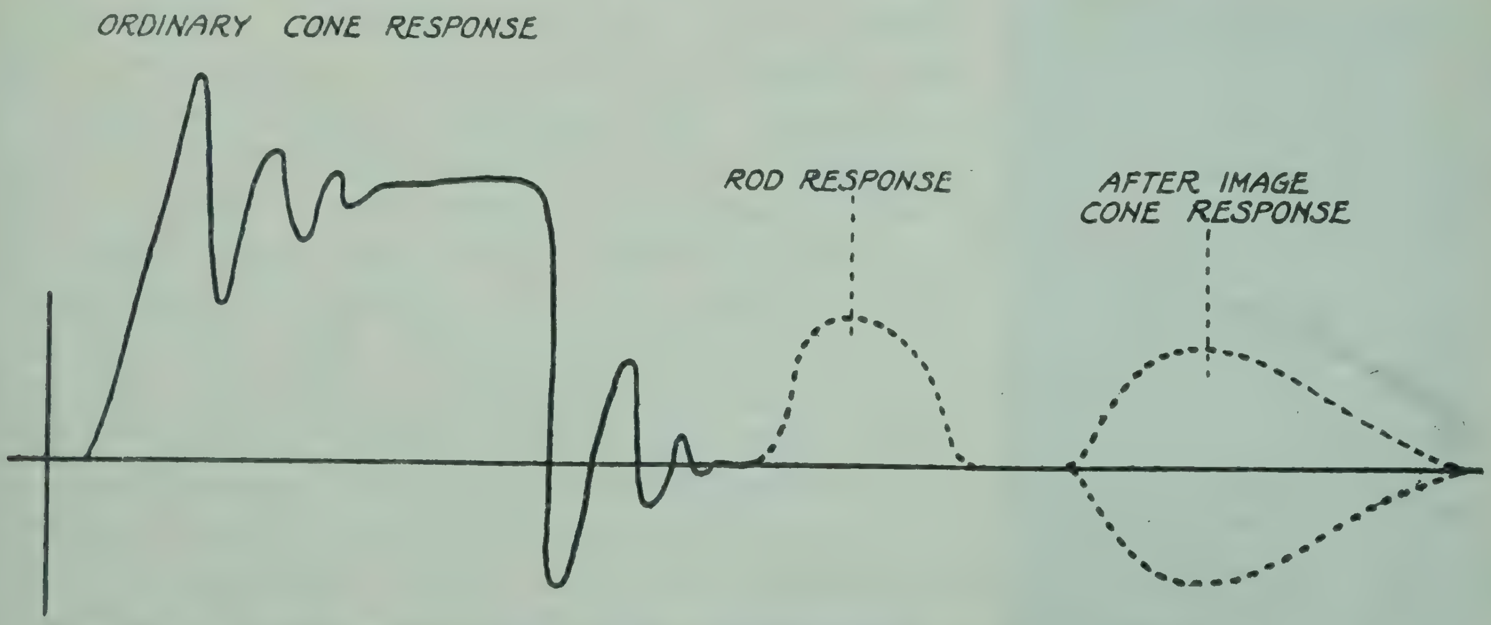


Fig. 286. Curve representing diagrammatically the Sensations aroused when the Eye has been Stimulated by a Flash of Light.
Intensity of Sensation vertical. Time horizontal. (Hartridge.)

be a shade of purple; but if nearer together than the complementaries, then the colour formed by the mixture corresponds to an intermediate part of the spectrum. Thus, if the colours mixed are red and green, the intermediate yellow and orange portions of the spectrum can be matched. As a rule the mixed colour is not so pure as the corresponding spectral colour, being less saturated (that is, diluted with a certain amount of white light). These facts can be expressed diagrammatically in the form of a geometrical figure, the colour triangle, in which the three fundamental colours occupy the corners and white the centre (see p. 373). Matches made by light of one intensity require readjustment if the intensity be changed, and matches made by one observer are different from those made by another.

For Reference

Hecht (1937). Rods, Cones and the Chemical Basis of Vision. *Physiol. Rev.*, **17**, 239.

8. THE SUBJECTIVE PHENOMENA OF VISION

By "subjective" we mean that the sensations under consideration cannot be traced solely to the stimulus which initiated them. Thus intermittent stimuli presented to the eye at a certain rate form a continuous sensation, so that flicker appears to have ceased. But the carrying on of the sensation from one stimulus to the next is performed by some part of the visual mechanism, and has nothing to do with any physical peculiarity of the light. It is, therefore, an example of a subjective phenomenon.

The Sensation Curve. When a light stimulus enters the eye a certain period of time elapses before a sensation is perceived. This latent period may be compared with that which occurs between the stimulus and contraction of a muscle. After its commencement the sensation rapidly rises to a maximum (see Fig. 286) and then shows several rapid fluctuations as it reaches its mean value. These fluctuations are called Charpentier's bands (Fig. 287), and are well seen after stimulating the eye by means of the flash from an electric spark. They have been compared with the oscillations which occur when an electric current is passed down a telegraph cable, and which are caused by the inductance and capacity of the circuit. Similar oscillations occur when the telegraph circuit is broken. Oscillations have also been witnessed and described by Bidwell at the end of the primary visual response. The time taken for the sensation to reach its maximum varies from 0·16 to 0·07 secs., being shorter the greater the intensity of the stimulus. This so-called primary image is followed under certain conditions by a less definite and less intense image, which has the following characteristics: (1) It is not seen when the eye is light adapted. (2) It is strongest for green light, and is absent for red. (3) It is absent from the fovea. (4) It is not seen by persons suffering from night blindness. (5) It is always of a bluish-grey colour. All the above facts fit in with the view that, whereas the primary response corresponds with the reaction of the cones to the stimulus, the image which is sometimes seen to follow belongs to the rod apparatus.

Fig. 287. Charpentier's Bands as seen when a disc with a narrow radial slit is rotated in front of an illuminated screen about once a second.

The After Image. Following these responses of the cone and rod end organs is the so-called secondary image, which certainly concerns the cones and may concern the rods as well. This image is of longer duration than those already considered, and it is of much lower intensity. It has, however, the peculiarity that, so long as it lasts, the part of the retina affected gives a diminished response to a stimulus of the same type as that which it had previously received. For example, if the first stimulus was one of white light, then a second white light stimulus falling in the period occupied by the after image of the first would not be recorded by the retina with the normal intensity. It has been pointed out that the after image period in some ways resembles the refractory period which follows the activity of muscle and nerve. If the second stimulus affects a larger area of the retina than that affected by the first stimulus, then a dark patch is seen lying on a bright ground, the dark patch corresponding in size to the first stimulus, the bright ground corresponding in size to the second stimulus. If the first stimulus is coloured, and no stimulus follows, the secondary image is found to have the same colour ; but if a second stimulus of the same colour, but rather larger size, falls on the retina during the secondary image, then, as before, a dark patch is seen lying on a bright coloured ground. If, on the other hand, the first stimulus be coloured light, say red, and the second stimulus be one of white light, the sensation received is one of the complementary colour, the reason being that the red constituents of the white light are partially excluded by the after image of the first stimulus, but not so the other spectral colours, and the image which is seen is therefore a blue-green one. Because of these peculiar properties the after image is said to have two phases : being called positive when the eye receives no second stimulus, and the appearance of the after image is the same as that of the first stimulus ; and being called negative when, owing to the incidence of a second stimulus, the after image shows the opposite intensity or colour to the stimulus which originated it. Since absence of the second stimulus causes the after image to be positive and the presence of a second stimulus makes it appear negative, we should expect a second stimulus of the right intensity to cause the after image to disappear altogether, since it would stimulate the surrounding retina with the same intensity as does the after image of the first stimulus. Experiment shows that this result can be achieved. Because of the importance of these properties of the after image, we may with advantage recapitulate as follows :—As a result of a stimulus the region of the retina affected gives a response which is followed by a second or after image. During this after image, this area is incapable of reacting with the normal intensity to a like stimulus, but shows increased excitability to a stimulus of the opposite kind. For example, after a green stimulus the retina is unable to respond fully to another green stimulus unless it falls either before or after the period of the after image. Therefore, if during that period a white stimulus be arranged to fall on the retina, it will cause a purple sensation (purple being white minus green) in that part of the retina first stimulated. The duration of the after image is variable, but is found to correspond roughly with the intensity and duration of the stimulus. Thus, after a few seconds' exposure to a bright light the after image may be noticeable for two or three minutes, its intensity waxing and waning in an irregular manner. Successive images are often found to show a series of colours ; a common series is bluish-green, violet, rose, and finally orange or green. The phenomenon is, however, very variable. The colours may be explained by assuming a difference in the rate of oscillation for the after images of the different colours. For example, the above series of colours would be obtained if green were more rapidly, and red less rapidly, perceived than blue.

Flicker and Visual Persistence. A study of the characteristics of the sensation curve provides an explanation of a number of the subjective phenomena of vision. For example, if a cardboard disc, marked as shown in Fig. 288, be caused to rotate slowly, while the black and white sectors are readily recognised, their radial margins appear blurred. This blurring is due to the slow rise and fall of the primary image of the sensation curve. If the speed be increased, a point is reached at which the disc gives an unpleasant glittering appearance ; this would appear to be due to one stimulus occurring during the after image of the previous one, and thus becoming suppressed, but being followed in its turn by a fresh stimulus which is caused by contrast (see later) to have a greatly increased intensity. If the rate of rotation of the disc be still further increased, a point is reached at which a stimulus falls during the primary image of the previous one. The persistence of the primary image after the cessation of the stimulus causes the stimuli to fuse to give a uniform sensation without flicker, which may be compared with the complete tetanus of a muscle. Since the primary response is more abrupt the greater the intensity of the stimulus, a more rapid rate of rotation is required to produce fusion at high intensities than at low.

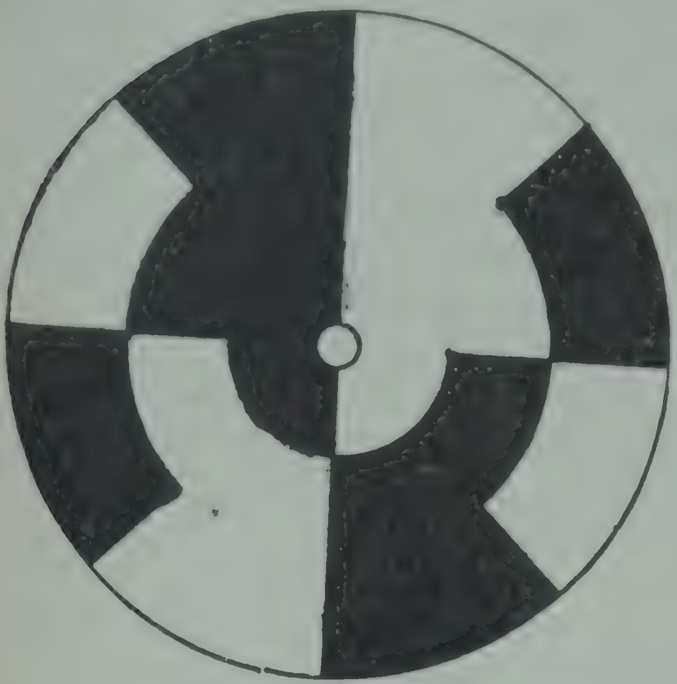


FIG. 288.

PERIODIC STIMULI. We have seen that, if a stimulus falls during the after image of a previous one, its character is altered. If the two stimuli are similar the second stimulus tends to be suppressed, but if dissimilar the second stimulus appears to be increased. If, on the other hand, the second stimulus falls either before or after the after image, it appears to be unaffected. But the experiments on flicker show us more than this, because even in a case where a second stimulus comes before the after image of the first, it is clear that the third would at a certain speed coincide with it, and should therefore be modified. Experiment seems to show no evidence for such an effect, and we must therefore conclude that the occurrence of the second stimulus in some way inhibits the after image of the first, so that its results are not apparent. Further evidence for this view is to be obtained from continuous stimuli, for we do not find a sudden diminution in the intensity of the response a moment or two after a continuous stimulus has begun, such as we should expect if the after image of the commencement of stimulus were suddenly after a short interval to assert itself. What happens to these suppressed after images? Are they entirely destroyed, or are they caused to accumulate until the end of the stimulus? The evidence appears to be in favour of the latter view, because an after image has a more definite character the longer the stimulus. Moreover if the gaze be directed towards a fixation point, and the inclination to blink be rigidly suppressed, after a few seconds the images of objects which fall on the periphery of the retina begin to appear milky, particularly in the shadows. At the same time the brightness of the high lights seems to be reduced so that it approximates more and more closely to the milkiness of the shadows. When this stage is reached, objects appear in outline, the contours being produced and renewed by imperfect fixation. If fixation can be retained for a short period, it will be found that the whole field becomes blank with the exception of the fixation mark. If this also disappears momentarily, then fixation is lost, the eye makes an involuntary movement, and the whole field immediately fills with detail again. In this experiment, two processes seem to be going on: firstly, in the shadows, the disappearance of the after images of previous impressions, the replacement of the visual purple previously bleached and possibly also the recovery of these parts of the retina from the effects of previous stimulation, all of which will increase the sensitiveness of the retina so that it now responds to the light reflected by the shadows; secondly, in the high lights, the accumulation of after images, the bleaching of the visual purple, and possibly the effects of fatigue, all of which tend to reduce the intensity of the impression. So that these processes, tending to increase the brightness in the shadows, and to decrease that of the high lights, finally bring them to the same level. In these processes the accumulation and removal of after images would appear to take a considerable part. The conclusion to which we are forced is, that at the beginning of a continuous stimulus, the after images are effectively removed until such time as the stimulus shall cease, when they can be permitted to assert themselves. But if the stimulus be prolonged, the suppression becomes more and more difficult, until the accumulation of after images is so extensive that they begin to obtrude more and more on the impression conveyed to consciousness.

Fate of After Images. If the conclusion drawn from the above experiments is valid, the question arises as to the apparent unimportance of the after image in ordinary vision. The answer is to be obtained from experiments like the following :—If fixation be continued until the images formed on the retina appear in outline as in the previous experiments, and the gaze be then quickly turned to a second fixation mark placed some distance from the first, it is found that, on returning to the first mark, some time has to elapse before the appearance in outline is obtained again ; in fact the time taken is not very different from that required to reach this stage at the beginning of a new experiment. The second impression had effected an almost complete removal of the after image of the first, so that on returning to the first again, the slate had, as it were, been wiped clean, and the first impression therefore acted as if it were a new one. This conclusion is entirely in agreement with our previous conclusions with regard to the after image, namely that it corresponds to a period in which a stimulus, similar to that to which the after image belongs, is inhibited, while that of a different kind is favoured. If, therefore, during fixation the gaze be momentarily directed elsewhere, the after image that had been set up is quenched by the new impression, and on returning the gaze to the fixation point the old image behaves almost like a new one. The non-intrusion of the after image in ordinary vision is therefore due to a considerable extent to the continual and rapid replacement of one impression by another, by the shifting of the gaze, and also to the spreading of the accumulation of partially effaced after images more or less uniformly over the retina. It has been suggested that impulses may be originated from the external eye muscles on movement, which on reaching the brain assist in the removal of the after images of previous stimulation.

Adaptation. If the eye after being in the dark is rapidly exposed to the light, at first the sight is confused and the eye dazzled, in spite of the powerful constriction of the pupils. The eye, however, very quickly becomes accustomed to the greater intensity, or, as we say, it becomes light adapted. In a similar manner, on entering the dark from the light, the eye can at first see nothing, but by degrees it becomes accustomed to the new conditions, and objects begin to be recognised : the eye has therefore become dark adapted. In the first case, the initial light stimulus reduced the sensitiveness of the eye to light to such an extent that the eye ceased to react to an excessive degree as it had done at first. Dark adaptation appears to consist of two separate processes : (1) the removal of after images from the cone light-receiving mechanism, and (2) the replacement of visual purple for the rod apparatus ; the former predominating at high intensities and the latter at low.

Fatigue of the Retina. If the eye has been exposed to a very bright light for a considerable time, there is at first inability to see with the dazzled part of the retina. If a field is looked at, a black spot appears to lie in front of it ; if, on the other hand, the field subsequently looked at is dark, this same area of the retina appears to be filled by a bright haze. If the dazzling light be restricted to one colour, there is an inability to see the same colour, if of lower intensity, immediately afterwards. The power to see other colours is apparently quite unaffected ; in fact Burch stated that the complementary colour actually appears to be more vivid than usual. These changes are similar to those caused by after images. The negative after image causes diminished appreciation of colours similar to itself, while the positive shows itself as a bright image similar in colour to the original stimulus when a dark field is looked at.

Contrast Phenomena. SUCCESSIVE CONTRAST. Visual impressions are affected by the previous history of the retina ; thus, after the eyes have been directed towards a red surface, a grey surface appears to be tinted green, a green surface seems a more vivid colour than normal, while a red colour is relatively dull. In other words, after stimulation by one kind of source, another of a similar nature is inhibited, while that of a different nature is either unaffected or may be even increased. This effect is called successive contrast. Experiment shows that the change of the second stimulus is such that it favours the colour complementary to the first stimulus. These effects are similar to those already described under adaptation and fatigue, and the causation of the phenomenon is the same as that given above, being due to the presence of an after image.

SIMULTANEOUS CONTRAST. If the retina is stimulated by two separate impressions, any differences between the impressions will be found to be accentuated. Thus if a small grey surface be placed on a white ground, the grey will become darker, and to a less extent the white ground will become lighter. If now the same grey surface be placed on a dark ground, it will be found to become lighter and the field darker. The nearer the surfaces are together the greater are the effects of contrast, the edges showing the effects of contrast most. In the case of colours, similar changes take place : thus two similar colours of different intensity placed together appear to be more different ; two colours of different saturation change, the one to greater, the other to less saturation ; while two colours of different wavelength appear, under the influence of contrast, to suffer a variation towards the complementary colour. A similar change is observed if contrast is occurring between a colour and a grey surface of approximately the same intensity, for we find that the grey is obviously tinted with a colour that is very nearly the complementary of the colour in question. It has been stated that the light which reaches the eye other than through the pupil (*e.g.* the sclera), and which is coloured an orange-pink in consequence of its partial absorption by the blood pigment in the capillaries, is responsible for the contrast colours not being strictly complementary to those which produce them. It is also found that separation of the surfaces, or the demarcation of the junction of the two surfaces by means of a narrow black or white line, or even the existence of small marks or creases, reduces the effects of contrast to a considerable extent.

We can summarise the effects of simultaneous contrast in the following way :—if one part of the retina is being stimulated, the part surrounding it not only tends to discourage a similar stimulus but also favours one of a complementary nature (black in this connection being considered as complementary to white). But this statement is similar to that which we have already made with regard to the effects of the after image on the same part of the retina which has received stimulation. Simultaneous contrast would therefore appear to be simply an extension of the after image phenomenon into a region of the retina outside the confines of the original stimulus. Experiment shows that this extension does not go far from the excited area, for the contrast effects, which may be considerable near the contour, rapidly decrease as the distance from the contour increases.

BIDWELL'S EXPERIMENT. If a brief white light stimulus be caused to fall during the after image period of a previous red stimulus, we should expect the white light to be tinted blue green, because red light is suppressed and its complementary increased. If the intensity and time intervals are carefully chosen, the blue green sensation may be made stronger than the original red stimulus to which it owes its colour. The persistence of vision causes this blue green response to last an appreciable time, and therefore if another red stimulus rapidly succeeds the white one (the one which is coloured blue green) it will tend to be suppressed from consciousness. But a white stimulus succeeding this second red one will be coloured blue green by its after image, because it has left its impression on the visual mechanism, although that impulse has not been conveyed to consciousness. If therefore a series of such red and white impulses be caused to fall on the retina, each white one will be tinted blue green, and each red one will be suppressed, the result being that the complementary colour alone is seen. This interesting phenomenon was discovered by Bidwell. He took a disc of tin plate about 8 inches in diameter, and arranged so that it could be rotated by an electric motor 6 to 8 times a second. From this disk a sector was cut of approximately 60 degrees; half the remainder was covered with black velvet and the other half with white paper. Behind the disc were mounted two pieces of silk: one red, the other blue green (Fig. 289). The order in which the images were presented to the eye on rotation of the disc was: (1) coloured silks, (2) white sector, (3) black sector, and so on. The result was found to be that the red silk appeared pale blue green, and the other pink; that is, in both cases the complementary colour alone was seen. This experiment brings out very clearly the fact that the after image process is entirely subconscious. The following observations confirm this conclusion.

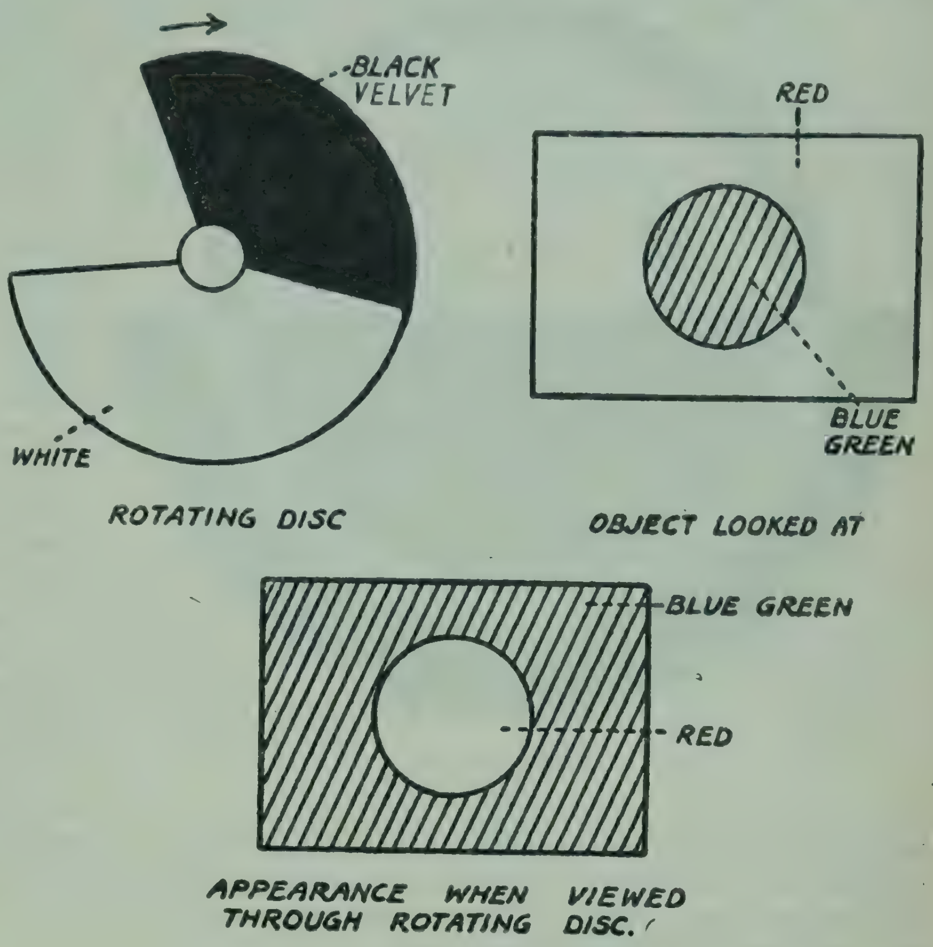


Fig. 289. Bidwell's Rotating Disk.

An object looked at through the disc appears in complementary colours.

Burch's Experiment. It has been much discussed in the past whether contrast phenomena are due to errors of judgment, as Helmholtz supposed, or due to physical or physiological changes taking place in the retina or in the conducting paths leading from it to consciousness. Burch disproved Helmholtz' view by the following simple experiment. A box (Fig. 290) is divided into two long compartments, *a b* and *c d*. At *a* the compartment is closed by a red glass-plate and at *c* by a blue glass-plate. Apertures are provided at *b* and *d* for the observer's eyes. At + and + two small grey crosses are fixed about the middle of the compartment on sheets of transparent glass. On looking through the openings *b* and *d* and converging the eyeballs, so as to fix the point *o*, we get a fusion, more or less complete, of the two colours red and blue, so that the background appears purple; or there may be a struggle between the colours, at one time blue, at another red predominating. To the judgment, however, there is one background and not two, and therefore, according to the theory of Helmholtz, the grey crosses should by contrast both acquire the same induced colour, which would be complementary for purple. But it is found that the two crosses are perfectly distinct in colour, that which is seen by the eye against the blue ground being yellow, while that on the red ground is green, showing that the phenomena of simultaneous contrast are not due to an error of judgment.

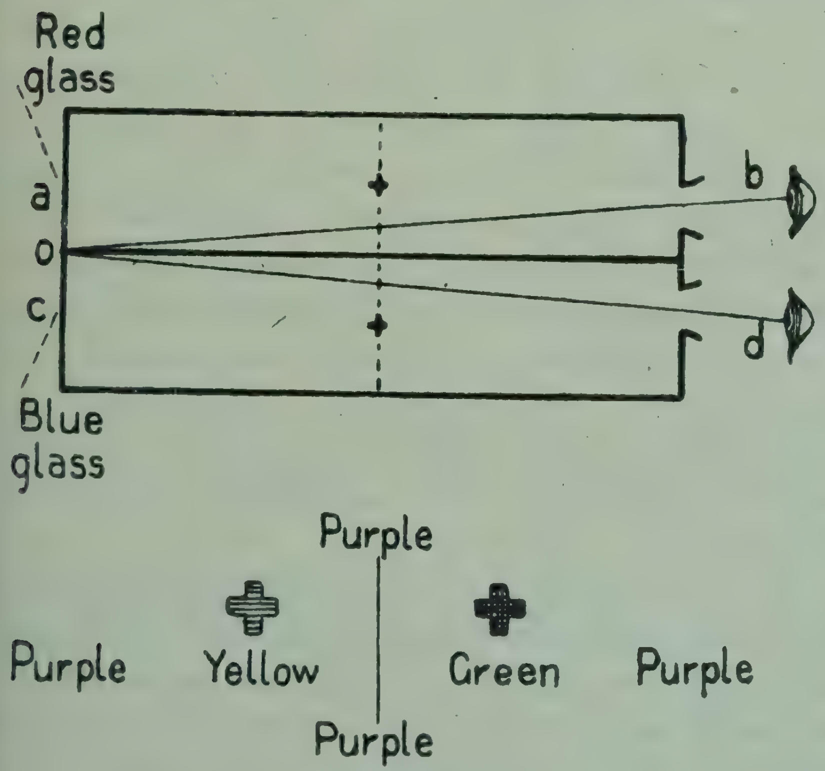


Fig. 290.

Sherrington's Experiment. The same fact is very definitely established by the following experiment devised by Sherrington. The disc (Fig. 291) presents two rings, each half-blue and half-black. The outer ring is intensified when at rest by simultaneous contrast, the black half being seen against the surrounding yellow, while the luminosity of the blue half is increased by the effect of the surrounding black. In the inner ring the blue half is darkened by contrast with the surrounding yellow, while the black half is not evident at all. If the disc be rotated, we get two concentric rings on an apparently homogeneous field. It is found however that the outer ring flickers long after complete fusion has taken place in the inner ring, showing that the stimulation of the retina by the outer ring is increased under the influence of contrast.

Allen's Monocular Experiments. By means of two independent monochromatic illuminators, Allen illuminated one eye with a small flickering light source of one colour lying on a background of a large non-flickering source of another. He varied the rate of flicker until a smooth sensation was perceived. He then made the justifiable assumption that, when an increase in flicker rate is required for smoothness, it demonstrates an increase in the sensitiveness of the retina at that point and *vice versâ*. Allen found that a red background caused a decreased sensitiveness of the retina to red, but an increased sensitiveness for green and blue. A green background on the other hand made green minus, but red and blue plus, while a monochromatic yellow background made both red and green minus, but blue plus plus. Now these results fit in exactly with older ones in simultaneous contrast. They show, however, more than these could do, namely that both facilitation and inhibition are at work. They show, moreover, a trichromatic subdivision of the spectrum such as Young's theory of colour vision postulates.

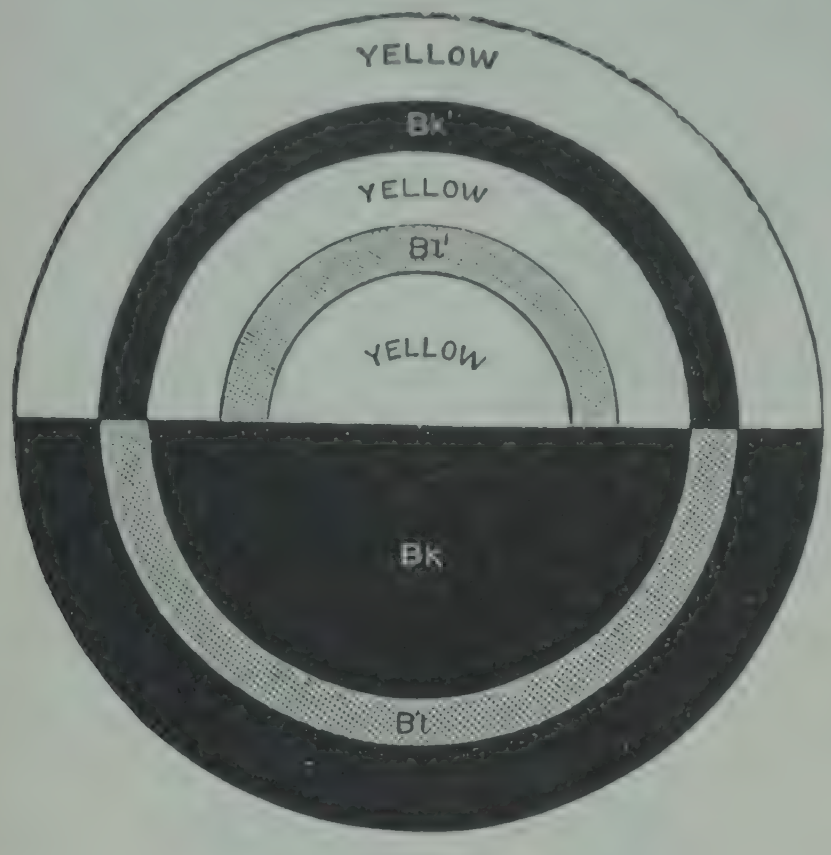


FIG. 291.

ALLEN'S BINOCULAR EXPERIMENTS. By a similar arrangement to that used above, the constant coloured field was presented to one eye, and the small flickering coloured source of light to the other eye. These experiments also demonstrate a trichromatic sub-division of the spectrum, since a stimulus of red, green or blue to one eye enhances the perception to the whole spectrum, but especially to the corresponding colour in the other eye; there being two invariable regions between these at 0·655 and 0·500μ respectively.

CAUSE OF AFTER IMAGE. Sherrington has shown that in the case of muscles there is what is called reciprocal innervation. Thus stimulation of the cortex which causes the contraction of one muscle also brings about a corresponding relaxation of its antagonist, in order that rapid and economical motion may take place. The contraction and corresponding relaxation are therefore analogous to the response of one part of the retina, which is accompanied by an inhibition of the surrounding parts of the retina to the same kind of stimulus (simultaneous contrast). In a similar manner the inclination to extension, which is found to accompany the prolonged flexion of a limb, finds its analogy in the phenomena of after images and adaptation of the eye, since the tendency is to suppress a similar stimulus in the part of the retina stimulated and to encourage its complementary. The inference to be drawn from these analogies is that the after image and its allied phenomena are caused by changes in the conducting paths of visual impressions which are similar to those found to exist in paths belonging to the motor system. What the nature of these changes in the conducting paths may be is at the present time undecided. McDougall has suggested that they are fatigue effects in the synapses of the higher conducting paths. If this view is correct, it would seem difficult to explain why after images and contrast phenomena are best seen with a rested eye.

UTILITY OF THE AFTER IMAGE. We have seen that the effect of the after image is to inhibit the possible reception of a second stimulus, and at the same time to favour the reception of one of a different nature. The process is therefore one which favours change, for not only is there a tendency to efface an old impression but also to welcome a new one. Such effects must be of great value to an organ such as the eye, the function of which to a considerable extent in everyday life is to present to consciousness the greatest number of impressions in a given time. For example, by measuring the time taken to read a passage in which almost every word was of importance, it was found that on an average eight words were read in each second, and that the eight words had an average of forty letters. It is clear from this that from eighty to forty different impressions must be presented to consciousness in each second. The function of the after image in preparing the retina for the reception of new impressions would therefore appear to be a very important one. The effects of simultaneous contrast are equally important to vision, because the changes produced by it are such that the images falling on contiguous portions of the retina are made as unlike as possible. Not only are the intensity and colour of adjacent parts of the image made more definite (this process being comparable to the effects of intensification on a photograph), but the blurring at the edges of contours due to imperfections in the image formed on the retina, is also largely eliminated (this comparing with retouching in photography).

Binocular Rivalry must be briefly referred to here, because of the similarity which it shows with the phenomena described above. If for any reason the images formed on the retinæ have dissimilar contours, rivalry ensues, first one image and then the other reaching consciousness. This process usually occurs independently in different parts of the field, so that the visual impression consists of a patchwork of the two images (Fig. 292). Seldom, if ever, are both images seen in the same part of the retina at the same time. It is found that a number of factors can cause one image totally to suppress the other : these are interest, novelty and brightness. The importance of this suppression can be appreciated by picturing the confusion which would occur if two different images were simultaneously presented to consciousness, as would happen in animals in which different images are formed in the two retinæ and in cases of strabismus in man. The parallelism between this process and those which we have already described can be traced by regarding for one moment one of the images as the primary one. This causes impulses to travel to the visual centres unaffected at first by the effects of other images, but these tend to accumulate more and more until the primary image is overcome, and that of the other eye put in its place. But this image in time suffers in the same way, so that the images alternate. The fact that a new image can suppress an old one, is due to the absence of after images in the first and their presence in the second. The predominance of a bright image can be explained by the longer time required for the after image to reach such a level as shall cause suppression. The preference for an image with contours would seem to be due to the greater ease with which the after image may be removed by small deflections of gaze.

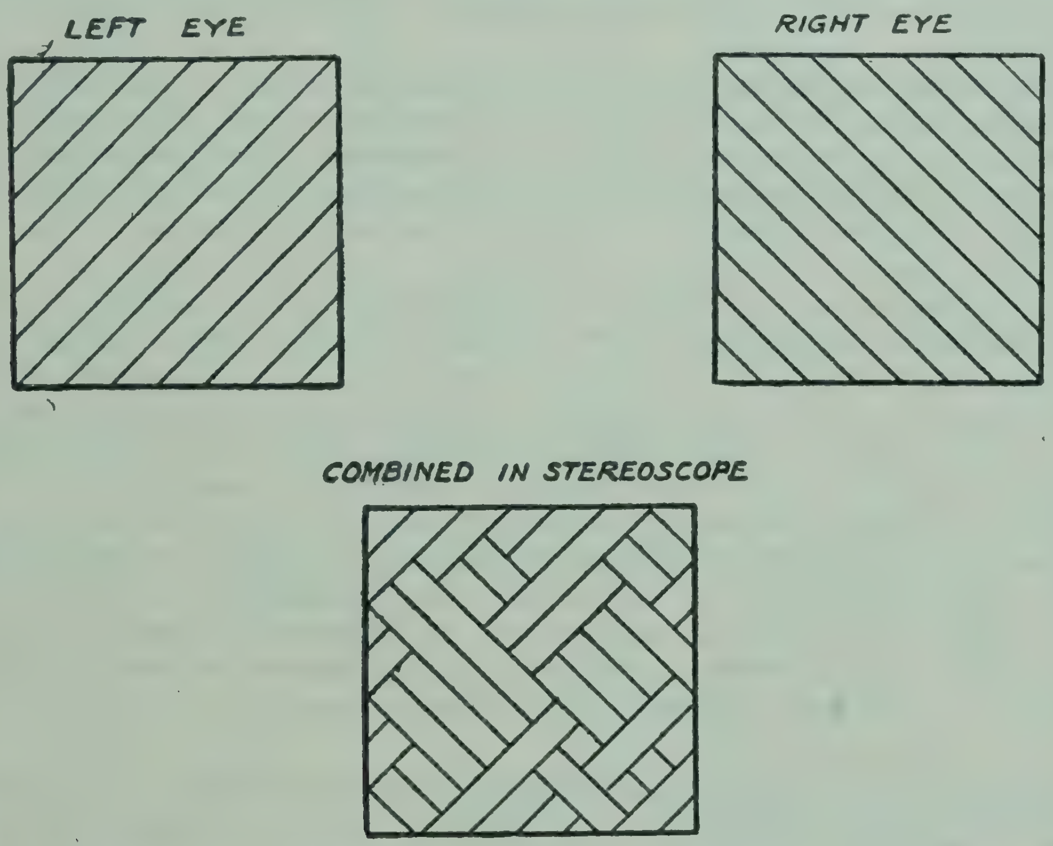


Fig. 292. Diagram to show how the Binocular Combination of two Dissimilar Images produces a Fluctuating Image containing parts of both of them.

9. ERRORS OF APPRECIATION

Under this heading we include all types of abnormality in the retinal apparatus or in its central connections. The class therefore includes cases n which the image formed on the retina is in every way normal, and those n which the optical defects do not adequately explain the whole of the visual disability which experiment shows to be present. The class is found to include ases which range from slight impairment to complete blindness.

The following classification may be used :—

Group 1. Both rod and cone vision are affected, and there is thus both night blindness and total colour blindness.

Group 2. Rod vision is either affected alone, or there is slight defect in cone vision as well.

Group 3. Rod vision is unaffected, while cone vision is either altogether absent, or is found to show abnormality which affects certain colours only.

Note that any one of the above groups may be found to affect either one or both eyes, and may involve the whole or only a limited part of the retina.

Complete Blindness (group 1) may affect the whole of one or both eyes, or may occur in half the visual fields only. It may be limited to irregular-shaped islands or patches, or it may be found associated with central or peripheral vision. The shape of the affected area frequently gives a direct clue to the cause of the condition. The shape is best determined by means of the perimeter.

(*a*) The whole of one or both eyes is found to be blind. The disease may be congenital or may be the result of inflammation affecting the posterior parts of the eye (ophthalmoscope will confirm). Injury involving the whole of the optic nerve trunk will also cause blindness of the corresponding eye.

(*b*) The blindness involves the right or left halves of one or both retinæ only. The lesion in such cases involves the optic tracts. Tumours are the commonest cause.

(*c*) The blindness is limited to a segment of the retina when the retinal vessels are affected (*e.g.* by embolism). The ophthalmoscope will confirm.

(*d*) Blindness which affects only the periphery of the retina is due either to deficiency of blood supply (as may occur in glaucoma), defective blood (severe anæmia), or the presence of poisons in the blood.

(*e*) If blindness affects chiefly the centre of the retina, the cause is probably poisoning by either tobacco, alcohol or both. At first vision is impaired for certain colours only but the blindness quickly becomes complete if the absorption of the poison continues.

(*f*) If irregular blind islands called *scotomata* are found in the visual fields, the cause may be inflammation of the choroid or of the retina itself, or the detachment of the retina from the choroid.

Night Blindness (group 2) is found in four different types of case.

(*a*) As an inherited condition. (*b*) In diseases of the liver in which bile salts are found to circulate in the blood, since these dissolve the visual purple out of the retina and therefore impair rod vision. (*c*) As a symptom of inadequate diet (lack of Vitamin A). (*d*) In disease of the retina or choroid (*e.g.* retinitis pigmentosa). In the last two cases it is usual to find colour vision affected to some extent.

The symptoms of night blindness are well described by the name. The eye does not possess the power of becoming fully dark adapted, and even a moderate degree is only attained after a prolonged stay in the dark. A photochromatic interval is not found, that is to say, when the intensity of a colour is reduced, it does not pass through an uncoloured stage (due to the rods). Purkinje's phenomenon is usually poorly developed or is not seen at all. There is no cure in the congenital cases.

Colour Blindness. The detection of this condition is important because of the use of coloured signals in the railway, the marine and other services. The employees of such services should be tested at frequent stated intervals, because colour blindness may develop (*e.g.* from tobacco or alcohol poisoning) in a relatively short space of time.

The methods of testing colour blindness fall roughly into three classes : (1) those of historic interest ; (2) those used by ophthalmologists for practical tests ; (3) those used for research. Holmgren's wool test is an example of the first class. It consists of a large series of coloured worsteds, the number of different shades being very great. In using the test the doctor hands to the patient selected skeins which have been found by experience to give difficulty to colour blind persons. He then instructs the patient to select from the box all the skeins that appear to him to have the same colour. His visual defect is judged by the mistakes which he makes. Since coloured wools readily fade and get dirty in use, slips of coloured glass or beads, etc., have been used in a similar way. These tests are however no longer employed, because parts of the test may so easily be lost or damaged.

Of practical methods the lantern test would appear to yield the best results, because it gives a close imitation of a signal light as seen at different distances and under various atmospheric conditions. In using the test, the examiner shows the subject a series of different coloured lights in turn, and in each case asks him to state what he sees. If he makes no mistakes, the colours are shown through modifying glasses which give the effect of a signal when seen through mist, fog, rain, etc., and the answers noted. One definitely wrong answer should reject the subject, particularly if red, green or white be one of the signals misnamed. Objections have been raised to this test, because it is possible for a man ignorant of colour names "to be failed" even if he has normal colour vision. The objection is readily met, however. When the driver of a train sees a signal, he says to himself "that is a green signal and therefore my train may proceed." But supposing all the time it were a red signal, and that he called it green through ignorance of the name, that man is as much a danger to the community as if he were in fact colour blind. No test can be too searching, and no borderline case should ever be passed; the risk is too serious.

Of tests of theoretical importance some have already been described, namely the measurement of the thresholds for light and colour, the colour-mixing apparatus and the flicker method of photometry. There is, however, another test which is found to give valuable information, namely the spectroscope test of Edridge Green. The instrument consists of a spectroscope to which is fitted two shutters, one of which may be caused to shut off the spectrum from the red end and the other from the violet. The patient commences the test by placing the shutter on the red side at the place where he sees the red begin, and this position on the wavelength scale of the shutter is noted. The patient is then told to move the other shutter until it reaches the place where red changes to orange. This wavelength also is noted. The red side shutter is now moved until it occupies the position of the violet side shutter, and the violet side shutter is now moved until a difference in colour at the two sides of the spectral area, which is thus isolated, is just not able to be seen. The wavelengths are again noted, and the next area measured off, and so on until the violet end of the spectrum is reached. A person with normal vision using this instrument will map out between 20 and 30 distinct strips. Abnormal vision may be shown in three ways: firstly, by the ends of the spectrum being found in abnormal positions, the spectrum being shortened at the red or the violet, or both; secondly, by the isolated strips being too large and too few; and thirdly, by wrong names being applied to some of them. The value of the method is considerable because it shows the presence of all three classes of defect, those due to blindness, those caused by ignorance of colour names and those in which the appreciation of colour is deficient.

Cases of colour blindness, when tested by the above methods, are found to show every possible variation between complete blindness and slight impairment of the colour sense. Their classification is as a rule complicated by the fact that cases are usually described in terms of one of the theories of colour

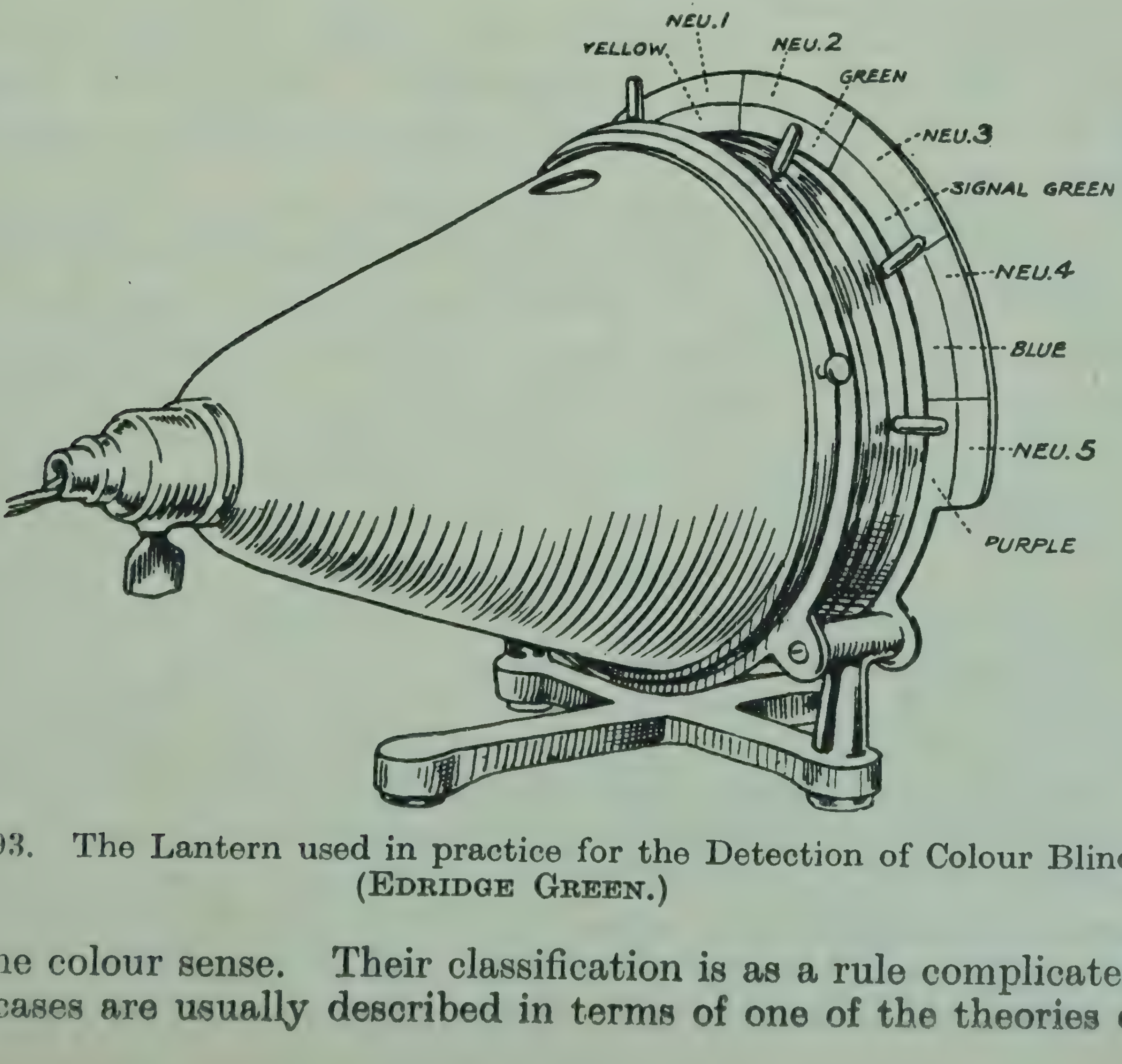


FIG. 293. The Lantern used in practice for the Detection of Colour Blindness. (EDRIDGE GREEN.)

vision. Most varieties of colour blindness are inherited, and are commoner in men than in women. But it may also be acquired, as explained above, in poisoning by alcohol and tobacco. Cases of colour blindness may be grouped as follows :—

(1) Cases in which the cone mechanism of the retina is not functioning. The patient is found to be colour and day blind ; red is not seen at all, while the other colours are seen as different shades of grey. Vision at night is good, vision by day is complicated by the fact that the patient must not expose his eyes to a bright light, for otherwise his visual purple will become bleached and his rod apparatus therefore cease to function. Owing to the absence of rods from the fovea, this part of the retina is blind. Visual acuity is therefore low. When tested by the flicker method, his luminosity curve is found to correspond with that of twilight vision. His condition may be improved by using neutral tinted glasses fitted with a sky shade.

(2) Cases in which the cone apparatus is apparently normal, that is to say, there is no avoidance on the part of the patient of strong light, no diminished visual acuity, no foveal blindness and no inability to see red light. Yet there is absolute inability to recognise all colours, any one part of the spectrum being able to be matched by any other. Tests by means of the flicker method show a luminosity curve which corresponds to that of day vision. It would seem clear that the retinal apparatus is in every way normal ; one is therefore forced to the conclusion that the defect concerns the brain centre which subserves the appreciation of colour. This view is supported by the fact that, between this extreme type and normal colour perception, there are a large number of cases which show various grades of defect. Some, for example, see two colours at the ends of the spectrum only, the intermediate portion being a neutral colour ; others see three only, at red, green and blue, and so on. Since in all these cases the cones are apparently normal, there would appear to be a parallel with cases in which there is no trace of deafness, and yet there is an inability to appreciate harmony, or to tell when two notes are in tune. In both types of cases, it would seem that the higher centres of perception and appreciation are either absent or are undeveloped. As might be anticipated, therefore, instruction and practice at colour naming and colour matching benefit a certain number of these cases, so that it is sometimes found that after such instruction the less abnormal cases are not readily detected. If, however, they are tested in a poor light, they are found to make mistakes which a person with normal vision would not commit. As these are the circumstances under which signals have frequently to be recognised, it is for this reason that the lantern test with its modifying glasses is so valuable.

(3) Cases in which certain parts of the spectrum are not seen at all. This condition frequently affects the red end of the spectrum, but it may be found in other parts. The principal effects to be noted are : (i) an abnormal type of luminosity curve, as examined by means of the flicker method ; (ii) the requirement of different amounts of the primary colours in order to match a given spectral colour, as compared with individuals of normal vision ; (iii) inability to recognise the normal number of different hues in the part of the spectrum affected, as shown by the spectroscope test ; (iv) shortening of the spectrum, if either of the ends of the spectrum is affected. These cases therefore show quite distinct features, which at once differentiate them from those of class 2. A typical example of a case belonging to class 3 will now be described, namely that in which there is shortening at the red end of the spectrum. The flicker test shows that the luminosity curve for different parts of the spectrum has its maximum in the green, instead of in the yellow. Further, the curve does not extend so far into the red, or show such high values in the orange as the normal curve. This curve therefore explains the apparent shortening of the spectrum. Colour mixture experiments show that the patient can match mixtures of green and blue with white light. When required to match yellow, he uses an excessive amount of red, and correspondingly less green than the

normal sighted; this test again shows the deficiency of the red sensation. Tested by the spectroscope, it is found that, besides the shortening at the red end of the spectrum, there is also an inability to distinguish the normal number of hues at the green and yellow. This effect is readily explained, because the difference between green and yellow shades largely depends on the varying extent to which they stimulate the red sensation. When the red sensation is absent, it is clear that the differentiation of greens and yellows must suffer.

10. THEORIES OF COLOUR VISION

The value of a theory to science is as much due to the fresh lines of research which it indicates, as to the explanation which it offers of the already ascertained facts. The theories of vision, therefore, are of value in spite of the fact that they do not at the present time offer a complete account of the retina and its functions.

THE DUPLEX THEORY of von Kries states that the retina shows two different types of response, one for low illumination and one for high, *i.e.* twilight vision, which is colour blind, and day vision, which responds to colour. The view that the rods with the visual purple supply the former, whereas the cones provide the latter, is one widely held and is already familiar, because it has been made the basis of the description given in previous sections. The evidence on which this opinion is based may with advantage be repeated because of its importance.

Twilight vision is found in those parts of the retina where there are rods; it is not present, therefore, at the fovea centralis, which consists entirely of cones. If a spectrum be examined, it is found that the colour with the greatest luminosity is the green, but that red rays are not seen at all. The form of the luminosity curve is identical with the bleaching curve of visual purple, and this pigment occurs only where there are rods. The acuity of twilight vision is low, and is explained by the fact that many rods as a rule send their impulses through one and the same nerve fibre.

Day vision is most highly developed in the fovea, from which rods are absent. Not only are the cones at the fovea placed very closely together, but it would appear that each cone connects to one nerve fibre only; in this way the high visual acuity is explained.

It has been suggested recently that the following modifications should be made in the duplicity theory of von Kries:—

1. That the cones do, in certain cases, function to some extent in night vision, thus retaining one of the primitive characteristics of rods, from which on morphological grounds they appear to have been developed.

2. That the fovea contains some visual purple, this being necessary in order that the cones may function in night vision as above, or possibly for the green sensation of vision.

3. That the rods play some part in day vision, adding their response to that of the day cones.

These modifications of the duplicity theory concern detail more than they do the basis of the theory, and do not appear to detract at all from the strength of its position. Hence, so far as the relative rôles of the rods and cones are concerned, there would not appear to be any room for speculation. Such is not the case, however, with regard to colour vision because, of the various hypotheses that have been so far advanced, none has been found to offer a feasible explanation of all the known facts, or to leave no other possible alternative. A brief account of the rival theories may with advantage be given.

YOUNG'S HYPOTHESIS states that there are cones of three different kinds, namely red, green and blue. The impulses from these cones are so combined in the brain that they give a complete picture of the separate coloured images. When all three types of cone are stimulated by white light, a colourless sensation results. Each visual unit may therefore be regarded as consisting of three cones, one of which responds to each of the fundamental colours. From this we should expect that the limit to the acuteness of vision would be reached when the separation of the images on the retina is not less than the diameter of such a unit. But the diameter of the foveal cones is approximately 0·0025 mm., and therefore that of a unit would be roughly 0·004 mm.

When the visual acuity of the eye is measured using monochromatic light it is found hat the limit is reached when the retinal images are separated by about 0·002 mm. The conclusion from this is, that the unit of response cannot be larger than one cone, and this uggests the idea that each of the cones of the fovea must be capable of responding to all hree of the funamental colours. There are thus two rival views for trichromatic vision : 1) Young's original hypothesis that there are three different kinds of cone; (2) the modified ypothesis that there is only one kind of cone, but that this is capable of responding to all hree fundamental sensations. These hypotheses will be considered in detail and then their elative advantages and disadvantages will be considered.

THE MULTIPLE CONE HYPOTHESIS. According to this hypothesis every cone f the retina is capable of responding to one colour sensation only. Thus in the fovea entralis there must be at least three different types of cone, those responding to red, those o green, and those to blue.

According to this hypothesis the brain receives nerve impulses which inform it of the colours which are present in the images formed on the retina. The possibility that every one of these cones behaves in an all-or-none manner was suggested by Hecht, but it is also possible that each of these cones responds in a graded manner sending impulses along the optic nerves which vary in frequency according to the intensity of light falling on the cone in question. This view, which was suggested by Hartridge, has been supported experimentally by Granit, by Hartline and others.

THE SINGLE-CONE HYPOTHESIS. According to this hypothesis every cone of the retina is capable of responding to more than one colour sensation, so that, for example, in the fovea centralis every cone can give responses to three colour sensations.

So far as is known no single nerve fibre is capable of transmitting more than one kind of nerve impulse. Moreover, each cone of the retina connects with one fibre of the optic nerve and no more. In consequence every cone must be capable of transmitting up its nerve fibre three different sensations, and this can only be done by variations either in frequency or in the pattern of the nerve impulses which are sent up to the brain. Suppose, for example, that when a cone responds to red rays it sends up its nerve fibre 100 impulses per second, whereas when responding to green it sends up 300, and for blue 1,000. Then provided that the brain has the necessary structure for recognising these three different frequencies, recognition of red, green or blue may be satisfactorily explained. When, however, two colour sensations are simultaneously being perceived there would have to be a transmission of two separate frequencies to the brain. It does not seem easy to picture the simultaneous transmission of two or more such frequencies.

Now since the type of colour sensation is determined by frequency, it is impossible for intensity of light to be transmitted by any variation in the number of nerve impulses per second sent up the particular nerve fibre. In consequence, it is a corollary of this hypothesis that each cone performs in an all-or-none manner and that different cones have different thresholds so that as the intensity of light increases the number of cones responding to that light increases, thus the active cone population will be small when light intensity is small and will increase to a maximum number as light intensity increases. Sooner or later, as this increase in light intensity continues the whole of the cone population will have been brought into a state of activity and no further increase of light intensity which may take place will cause a corresponding increase in the intensity of the sensation received by the brain.

EVIDENCE AGAINST THE SINGLE-CONE HYPOTHESIS. Granit found irrefutable evidence against the single-cone hypothesis when he discovered that a micro-electrode applied to the different fibres of the optic nerve of an anæsthetised or decerebrate animal picked up differences of electrical potential, when the retina as a whole was exposed to light. Thus one fibre might be connected to a receptor with a marked green response while another fibre might be connected to a receptor with a markedly orange response. Thus in the frog's retina there were found to be one type of receptor with a broad response, which he called a dominator, and five different types of receptors with responses in narrow spectral regions, which he called "modulators." (See Table on p. 440).

It will be shown later that we have strong evidence for postulating a similar arrangement in man's retina. Now this idea meets with difficulties concerning the acuity of the eye by white light and by coloured lights, because it would be expected that acuity for fine detail would be superior when white light is in use to that found when coloured lights are in use instead, because white light acts as a stimulus for all receptors irrespective of colour property, whereas coloured lights act as stimuli for only a few. Thus if the acuity by white light be made equal to unity, that by red or by green light might be expected to have a value of one-half, or perhaps one-third. But experiment demonstrates no such large differences, as is clearly shown in the table below.

RELATIVE ACUITIES BY LIGHTS OF DIFFERENT COLOUR

Observer	Test object	White	Red	Green	Blue
Shlaer *et al*	Grating	1	0·91	1·02	1·00
,,	Landolt's C	1	1·00	—	1·00
Hartridge	Letters	1	0·82	0·96	0·90
,,	Landolt's C	1	0·84	1·00	0·92
	Mean values	1	0·89	0·99	0·92

Thus acuity, when using red light or blue light, appears to be slightly less than that obtained when using white light or green light. But in fact the differences are within the experimental error of the method.

We see then that expectations from the multicone hypothesis, that acuity by white light would be superior to that by coloured lights, is not fulfilled. How can this discrepancy be explained ? An explanation can be obtained on the basis of the cluster hypothesis.

THE CLUSTER HYPOTHESIS. In any random distribution of several different populations there are always places to be found where one population predominates over the others. An excellent example of this is the colour screen of the Lumiere photographic plate. In this there are distributed irregularly three different coloured groups of starch grains, red grains, green grains and blue grains. When such a screen is looked at with the naked eye, red, green and blue dots are clearly visible. On examining these dots with a microscope it is seen that they are not due to isolated starch grains but to quite large groups of grains of almost entirely one colour. Now if the cones of different colour-response are distributed irregularly in the same way, we should expect to find places where the red cones predominate, other places where green cones are in excess, and other places again where blue cones are nearly exclusively present, and so on. In consequence, when high visual acuity is wanted for any one colour it might be obtained if it were possible by slightly rotating the eyes, to cause the image of the object in question to fall on a part of the retina where the cones predominate, which are sensitive to that particular colour. In the case of yellow light, a spot where there is a predominance of both red and green cones with the almost complete exclusion of blue cones, would be most suitable for the purpose.

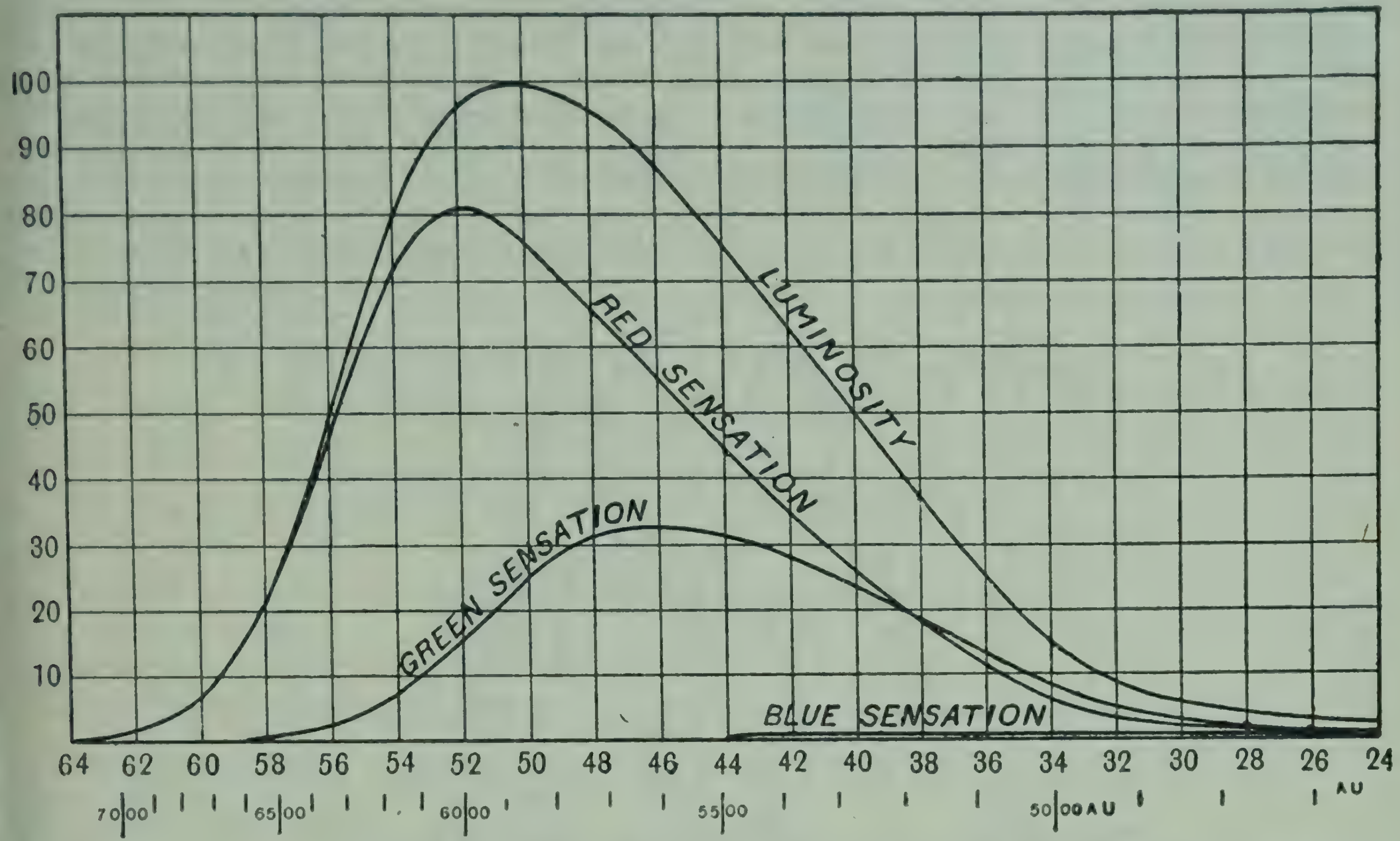


FIG. 294. The red, green and blue Sensation Curves and the Luminosity Curve of white Light. Luminosity vertical, wavelengths horizontal. (ABNEY.)

THE POLYCHROMATIC THEORY. Three independent lines of research give support to the view that human vision is polychromatic and not trichromatic as has long

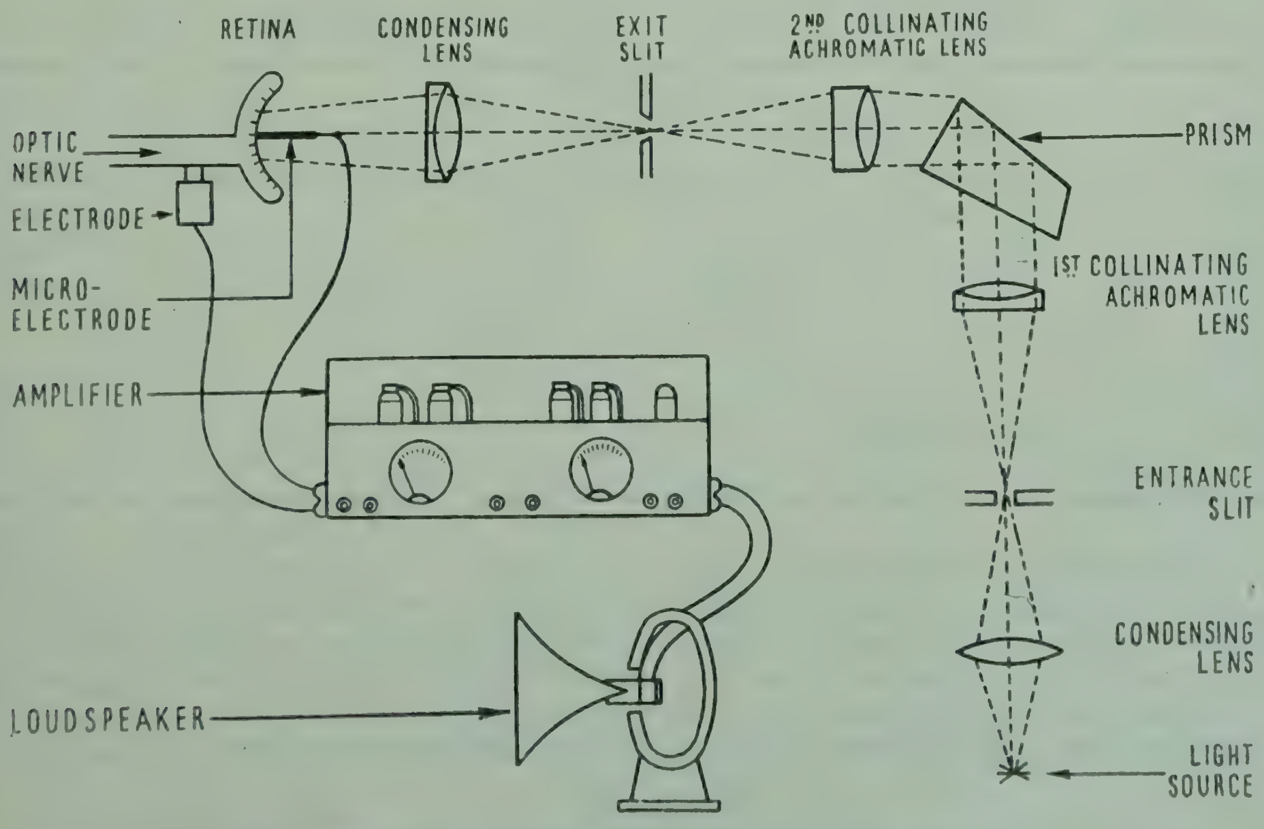

RETINA
CONDENSING LENS
EXIT SLIT
2ND COLLINATING ACHROMATIC LENS
PRISM
OPTIC NERVE
ELECTRODE
MICRO-ELECTRODE
1ST COLLINATING ACHROMATIC LENS
AMPLIFIER
ENTRANCE SLIT
CONDENSING LENS
LOUDSPEAKER
LIGHT SOURCE

FIG. 295. Granit's Micro-electrode Apparatus.

been supposed, and that the three-colour theory is an approximation which is applicable solely to foveal colour vision at medium light intensities and visual angles. The three lines of research are: (1) the micro-electrode experiments on animals, of Ragnar Granit; (2) the biochemical experiments on vitamin A derivatives, of Morton and his associates; and (3) the human experiments of Hartridge.

Granit's technique, which is illustrated diagrammatically in Fig. 295, consists in removing the anterior eye structures from a deeply anæsthetised or decerebrate animal and in applying to the fibres of the optic nerve a fine wire electrode which is coupled to a valve amplifier of high gain, the output of which is either connected to a loudspeaker or to a cathode ray tube according to requirement. The retina is illuminated by means of a monochromator, the experimenter's object usually being to determine for the different parts of the spectrum the minimum light energy required in order to produce a noticeable response on the part of the optic nerve fibre with which the electrode is in contact. After the examination of one nerve fibre the electrode is raised and is placed into contact with another, and so on.

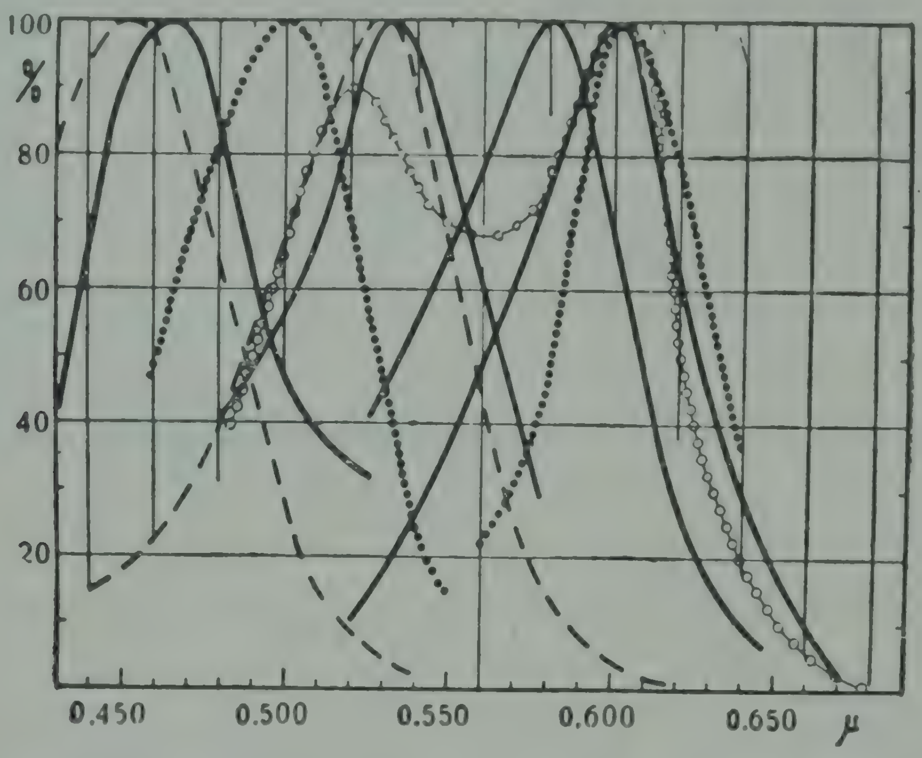


FIG. 296. Granit's Modulator. From "The Sensory Mechanisms of the Retina," 1947.

Granit has discovered in this way two essentially different kinds of receptors which he has called "dominators" and "modulators." The former exhibit wide response curves which extend over the greater part of the visible spectrum. The latter exhibit narrow response curves which are restricted to a particular spectral region; see Fig. 296. The positions of the crests of the latter in the spectrum are shown in the table below.

THE MODULATORS FOUND IN DIFFERENT ANIMALS (GRANIT)

Wavelength	Cat	Guinea-pig	Rat	Snake	Frog
0·600	Found	Found	Found	Found	Found
0·580	Found				Found
0·540	Found				Found
0·530		Found		Found	Found
0·500		Found	Found		?
0·480					
0·460	Found	Found			Found
0·420	Found at 0·440				

Ball, Collins, Morton and Stubbs (1948),* on applying strong acids such as concentrated sulphuric acid and syrupy phosphoric acid to vitamin A_1 or to retinene$_1$, which is itself an aldehyde of the former, find that they obtain a number of pigmented substances which have narrow modulator-like absorption bands which occupy characteristic positions in the spectrum. Moreover these pigments resemble visual purple in being sensitive to light. The positions of the crests of their absorption bands are shown in the following table.

---

* BALL, COLLINS, MORTON and STUBBS. *Nature*, 1948, **161**, 424.

λ	Frog modulators	Vit. A (H_2SO_4)	Vit. A. (H_3PO_4)	Retinene (H_2SO_4) 1	Retinene (H_3PO_4) 2
0·66μ	—	—	—	0·664μ	—
0·60μ	Found	0·62μ	0·60μ	—	—
0·58μ	Found	0·59μ	—	—	0·59μ
0·54μ	Found	—	—	0·56μ	0·55μ
0·52/53μ	Found	0·53μ	0·52μ	0·52μ	—
0·50μ	— *	—	—	—	0·50μ
0·48μ	Found	0·465μ	0·44μ 0·48μ	0·44μ 0·46μ	0·47μ
0·42μ	— †	—	—	—	—

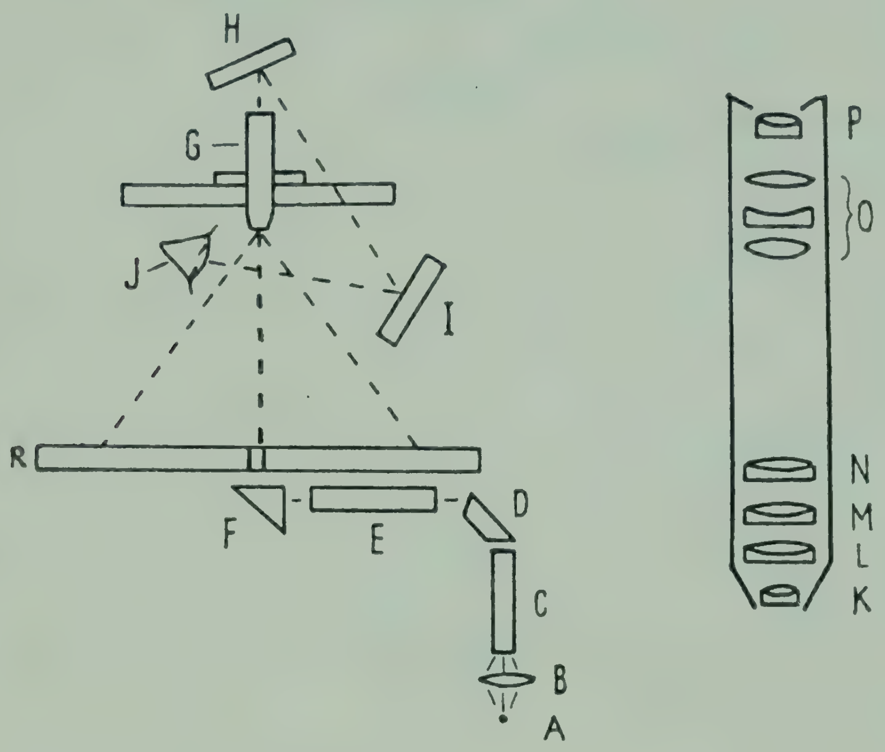


FIG. 297. Hartridge's Microstimulation Apparatus.

Light from the source A is focussed by the lens B into the monochromator C, D and E. The coloured rays emerging from the latter are reflected vertically upwards by the prism F. These coloured rays, together with rays from other sources placed on the board R, are greatly reduced in visual angle by means of the optical system G, some details of which are shown on the right at K to P. The reduced images are viewed by the eye of the observer J via the two plane mirrors H and I.

Hartridge, using a microstimulator, which presents to the observer's eye one or more minute points of coloured or monochromatic light (see Fig. 297), finds evidence for modulator-like receptors in the human retina, occupying certain characteristic spectral regions. These are shown in the table below.

Wave-length	Animal modulator	Human fixation point	Human subjective colour	Receptor	Unit	Wave-length
0·66		0·66	0·64	Red . .	R–BG–R	0·660 ?
0·60	Cat, Guinea-pig, Frog	0·62	0·61	Orange .	Tricolour	0·600
0·58	Cat, Frog	0·58	0·575	Yellow .	Y–B	0·578
0·54	Cat, Frog	0·54	0·54	Green .	Tricolour	0·547
0·53	Guinea-pig, Frog	0·52 ?	—	—	—	—
0·50	Guinea-pig	0·50	0·51	Blue-Green	R–BG–R	0·500
0·48			0·48	Blue. .	Y–B	0·476
0·46	Cat, Guinea-pig, Frog	0·46	0·45	Indigo .	Tricolour	0·450
0·42	Cat (at 0·44)	0·44		Red . .	R–BG–R	0·420 ?

* Found in the guinea-pig and the rat.
† Found in the cat, at 0·44μ.

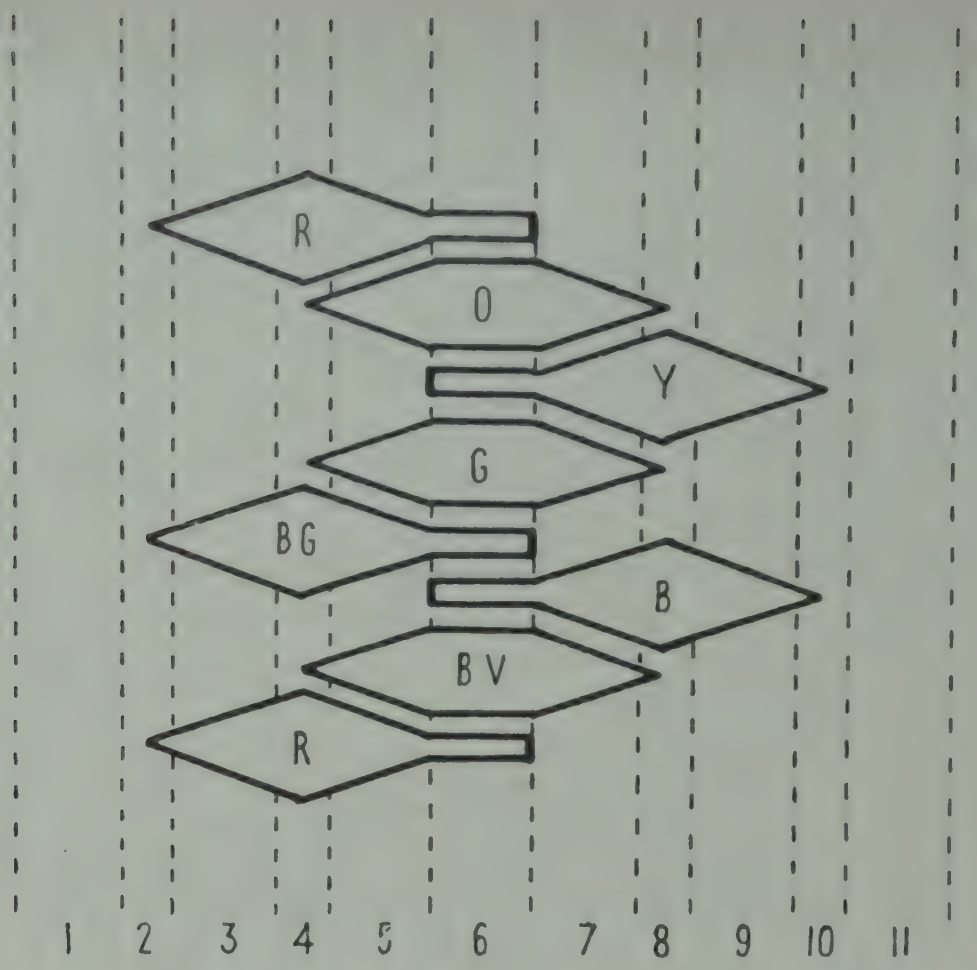


FIG. 298. The areas R, O, Y, G, BG, B, and BV indicate approximately, by their widths, the degree of activity of the receptors for red, for orange, for yellow, for green, for blue green, for blue and for blue violet, under different conditions of vision. (1) For the fovea at very low light intensities or very small visual angles. (4) For moderately low light intensities and rather small visual angles. (6) For the fovea at ordinary light intensities and visual angles of two degrees. (8) At the fovea for high light intensities, and in the near periphery of the retina for ordinary light intensities. (11) At the fovea for excessively high light intensities and at the extreme periphery of the retina for ordinary light intensities.

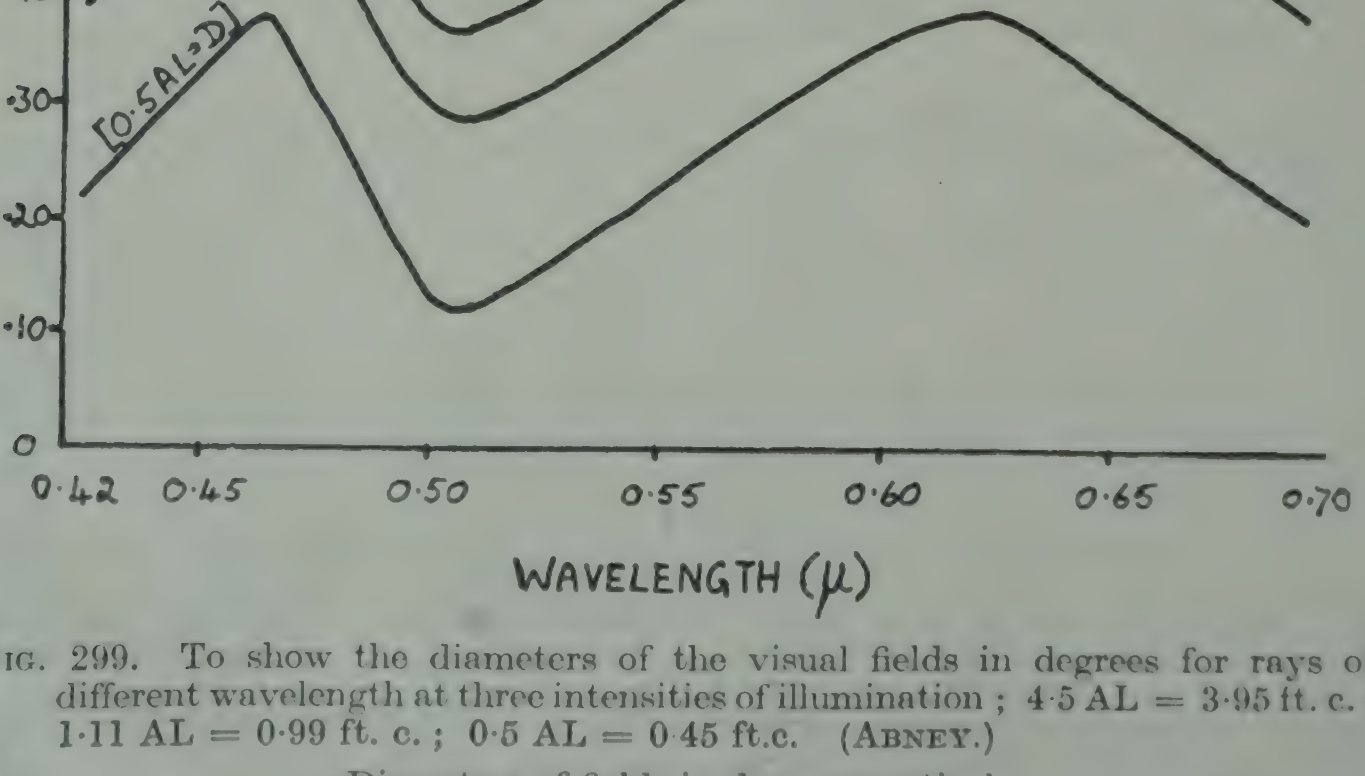


FIG. 299. To show the diameters of the visual fields in degrees for rays of different wavelength at three intensities of illumination; 4·5 AL = 3·95 ft. c.; 1·11 AL = 0·99 ft. c.; 0·5 AL = 0·45 ft.c. (ABNEY.)

Diameters of fields in degrees vertical.
Wavelengths (μ) horizontal.

Thus there are seven varieties of receptor in man, six of which have simple modulator-like response curves resembling those found in animals (see Fig. 296); the seventh one has twin responses, one in the extreme red and one in the extreme violet. If this receptor by itself was stimulated by light it would produce a sensation of a crimson or mauve light.

According to the polychromatic theory these receptors are arranged in three units: (*a*) a tricolour unit comprising the orange (0·600μ), the green (0·547μ), and the indigo (0·450μ) receptors; (*b*) a dichromatic Y–B unit comprising the yellow (0·578μ) and the blue (0·476μ) receptors, which are complementary in colour to one another; and (*c*) a dichromatic R–BG–R unit comprising the red (0·66μ and 0·42μ) and the blue green (0·50μ) receptors. These also are complementary in colour to one another. At the normal fovea, at medium light intensities and average visual angle, the tricolour unit has the greatest activity, but it is assisted by both the other units; see Fig. 298 at "6."

At high intensities of illumination the fovea exhibits a yellow-and-blue dichromatism due to the Y–B unit, the other two units having become inactive; see Fig. 298 at 9. At higher intensities still all colour vision disappears and there is colourless foveal vision; see Fig. 298 at 10 and at 11.

At low intensities of illumination the fovea exhibits a crimson-blue green dichromatism due to the R–BG–R unit, the other units having gone out of action; see Fig. 298 at 4. At lower light intensities still, all colour vision ceases and there is colourless foveal vision; see Fig. 298 at 2 and at 1.

Peripheral colour vision bears strong resemblances to foveal vision at high intensities of illumination; that is, there is a yellow-blue dichromatism at medium perimetric angles, and colourless peripheral vision at large ones. In Fig. 299 are shown the sizes of the perimetric fields as measured by Abney at four colour brightnesses. It will be noticed that the blue-green and the red and violet ends of the spectrum are the colours with the smallest fields, due to the R–BG–R unit having a small field. On the other hand, the yellow and blue regions of the spectrum have the largest fields, due to the large field of the Y–B unit.

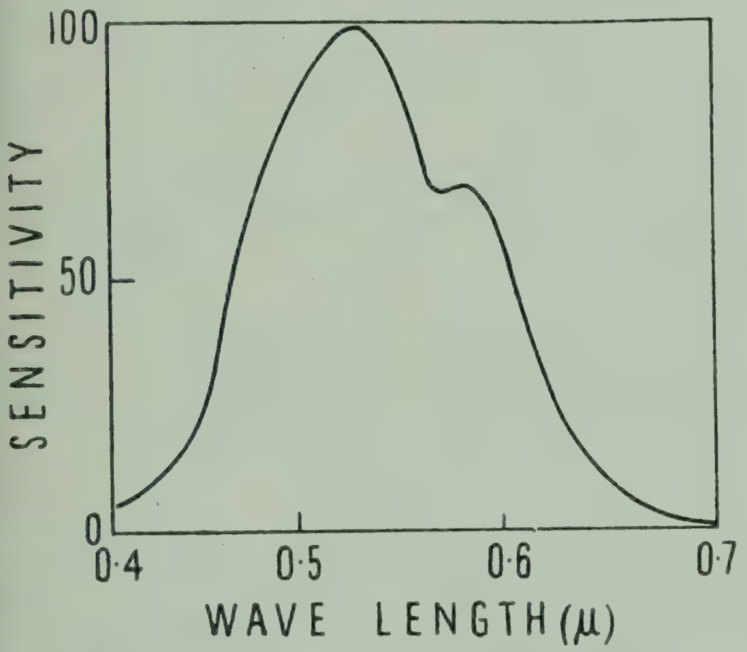


FIG. 300. Wright's notch in the foveal luminosity curve in the yellow part of the spectrum due to the use of a test object subtending a small visual angle at the observer's eye.

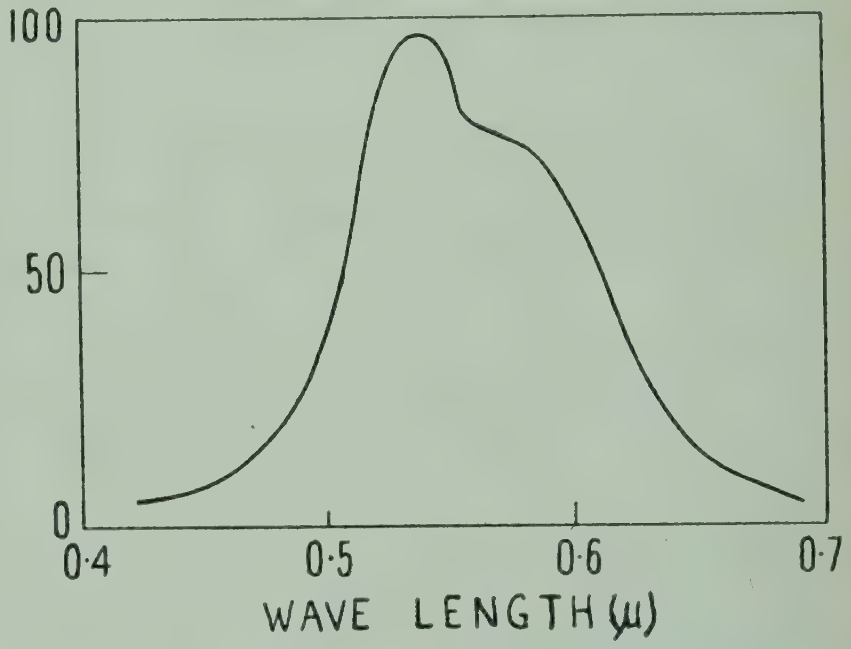


FIG. 301. Sloan's notch in the foveal luminosity curve in the yellow part of the spectrum due to the use of a low intensity of illumination.

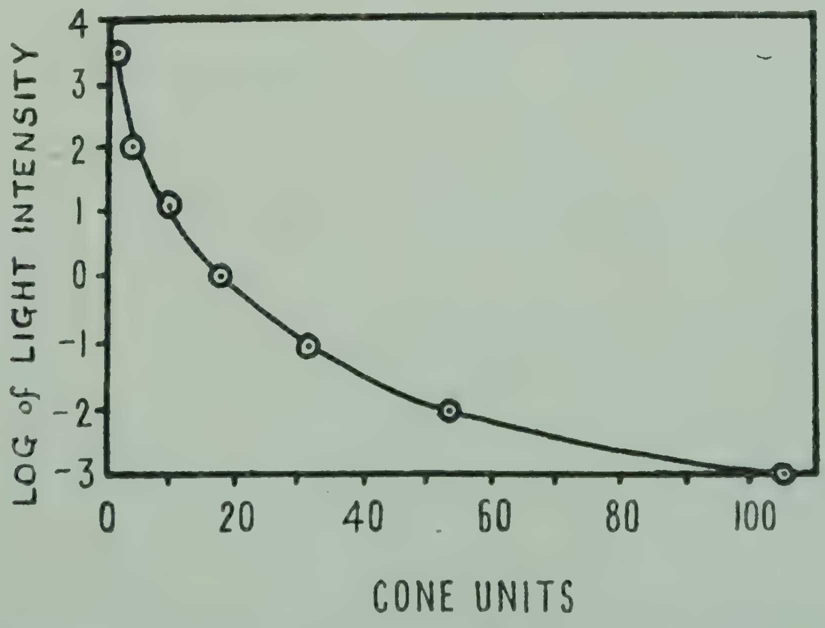


FIG. 302. Relation between light intensity and visual angle subtended by test object (see text).
Log of light intensity, vertical. Visual angle, in cone units, horizontal.

At small visual angles, the fovea exhibits a type of vision which is similar to, if it is not identical with, that found at low light intensities. Here are some points in favour of this view : (*a*) Both cause in an early stage the replacement of yellow by pale grey or white, and of blue by dark grey or black, and at the same time notches develop in the luminosity curve : Sloan's notch with reduction of light intensity ; see Fig. 301 ; and Wright's notch with reduction of visual angle ; see Fig. 300. (*b*) The visual angle at which the change takes place, either of yellow to white or of blue to dark grey, is altered by changes in light intensity. A high intensity causes a smaller visual angle to be required, and *vice versá*, as shown in Fig. 302. (*c*) With sufficient reduction of visual angle, or light intensity, or both, a red-blue-green dichromatism develops. And lastly, (*d*) with a still further reduction of light or angle, or both, all colour vision is lost and there is colourless foveal vision. In these and other respects the two changes in foveal vision resemble one another closely, if they are not identical.

With regard to the red-blue green dichromatism mentioned above, Fig. 303 shows the spectral coefficient curves obtained by Willmer and Wright for this type of foveal vision.

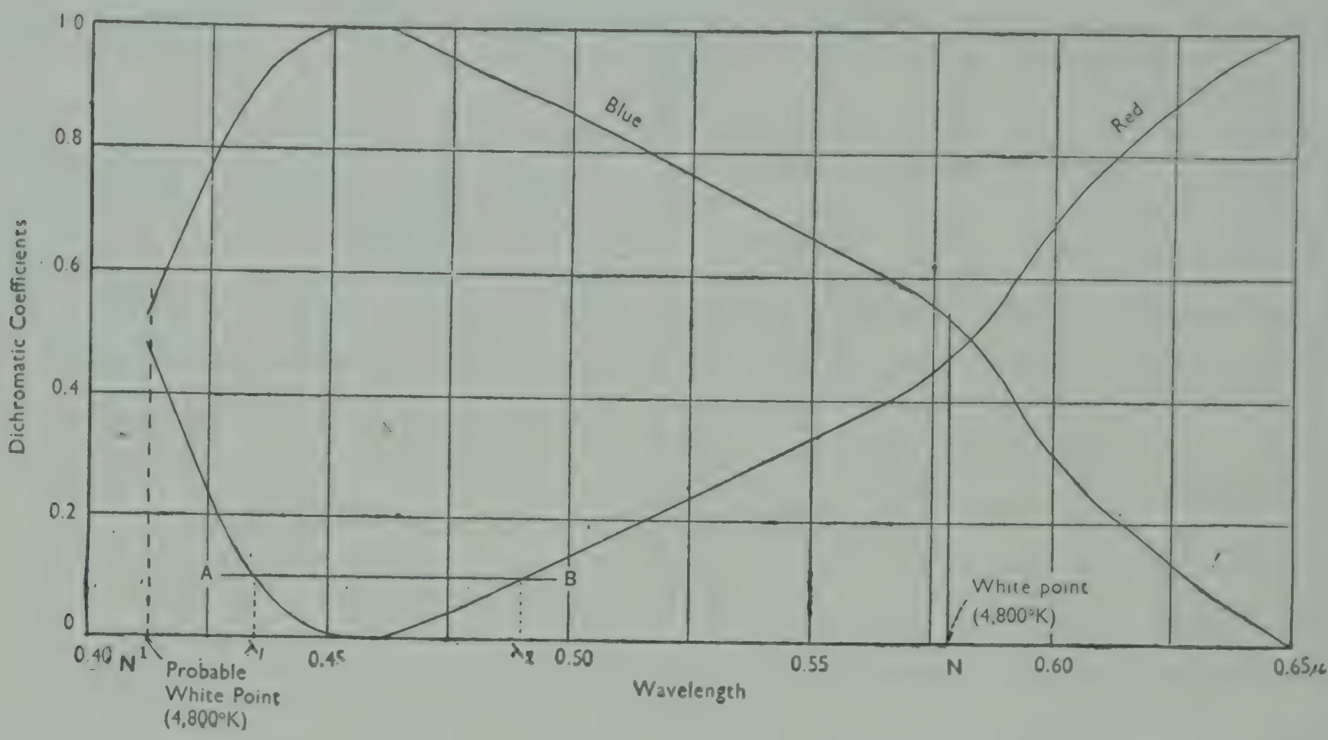


FIG. 303. Spectral coefficient curves for the fovea using 15 minute fields. Two neutral points are found ; one in the yellow at N and one in the blue at N¹. Arbitrary units vertical. Wavelengths in μ horizontal. (WILLMER AND WRIGHT.)

It will be observed that the red curve which falls to a low value at the medium wavelengths of the spectrum rises towards both the long and the short wavelengths. This is typical of the R receptors of the R–BG–R unit. The green curve gives its maximum response at medium wavelengths and decreases in activity toward either end of the spectrum. There are two neutral points in the spectrum, one in the yellow, the other in the blue, which are marked by N and N′ in Fig. 303. It will be recollected that yellow and blue are the two colours which are replaced by grey when either visual angle or light intensity are reduced.

THE CORRECTION OF THE CHROMATIC ABERRATION OF THE EYE. It will be remembered that the chromatic aberration of a lens causes the production of narrow yellow and blue fringes, which surround large objects and which replace small ones. The same statement applies to the lens-system of the eye. It might be supposed, therefore, that the correction of this aberration of the eye by means of a Helmholtz lens consisting of a combination of crown and flint lenses, or its elimination by means of monochromatic light, for example sodium light, would bring about the removal of these fringes and would therefore improve the acuity of the eye. But experiment shows no improvement, and moreover the fringes are invisible when the eye is used for normal vision. The causes of these apparently contradictory statements have long been under discussion. The true explanation is as follows : In the case of a test object constituted of a white dot on a black background, chromatic aberration causes the retinal image to be a yellow dot surrounded by a blue halo. But owing to the small visual angles which both dot and halo subtend at the eye, yellow is replaced by white and blue by black, so that the original appearance is returned to. In the case of a black line on a white background, chromatic aberration substitutes a blue line surrounded by a yellow halo on a white background. Now, as before, yellow becomes white and blue becomes black, and the original appearance is restored. By means of lens systems of special design it is possible to increase the chromatic aberration of the eye. When this is done it is found that the correcting effect present in the case of the eyes of one observer is sufficient to correct about double the natural aberration present. This is the probable reason for being able to use simple convex lenses, uncorrected for colour, in conjunction with the eye, as spectacles and as magnifying glasses.

The Colour Responses of the Human Eye. We have seen that the colour fringes which are introduced into the retinal image by the aberrations of the lens system of the eye are removed again by a process which consists essentially of an alteration in colour vision. Thus yellow and blue test objects viewed at large visual angles are seen in the proper hues, but the yellow and blue fringes, being viewed at small visual angles, are not. How are these changes in colour vision brought about ? Hartridge investigated a number of possible mechanisms and decided that the only feasible one was neurological in nature, that is, a nervous process alters the normal polychromatic vision of the fovea to a red-blue-green dichromatism. Now it is obvious that identical questions arise with regard to the other changes in foveal colour vision which occur, for example, with change of light intensity or of adaptation. Moreover at the periphery of the retina alterations of light intensity or of visual angle also alter the type of colour vision met with. How are these brought about ? Are they also caused by some neurological " response " ? Only further research can supply us with the answers to these and related questions.

For Reference

Wright (1946). Researches on Normal and Defective Colour Vision. (London.)
Granit (1947). Sensory Mechanisms of the Retina. (Oxford.)
Hartridge (1947). The Visual Perception of Fine Detail. *Phil. Trans. B.*, **232,** 519–671.

11. BINOCULAR VISION

Binocular vision may be defined as the co-ordinated employment of two separate visual organs in order to produce a single mental impression. The advantages of binocular as opposed to monocular vision are :—

(1) Optical defects of one eye are less important, since they are masked by the well-defined images of the other eye.

(2) Defective vision in parts of the visual fields of both eyes is hidden, so long as the defects do not affect the same parts of both fields. Thus under ordinary circumstances, each blind spot does not obtrude itself because in the other eye the image of the same part of the field falls on functional retina.

(3) The combined fields of the two eyes are larger than either alone because, while the features restrict the nasal halves of the fields of both eyes, the combined field contains the unrestricted temporal areas of both retinæ.

(4) Binocular vision under certain circumstances provides a very accurate perception of depth, size and distance, which is called stereoscopic vision.

In order that there should be binocular vision the following conditions should be complied with :—

(1) The fields of the two eyes must overlap. Animals in which the eye axes are parallel have the greatest overlap, and therefore possess the most perfect binocular vision.

(2) Approximately similar images must be formed on the retinæ, because if this condition is not satisfied, antagonism between the images will occur, as described above, and first one image and then the other will be presented to consciousness.

(3) The retinæ must possess physiologically corresponding points, in order that similar images formed on them may produce one conscious impression.

(4) The external eye muscles must so adjust the visual axes, that the centres of the fields of the two eyes coincide with the images of one and the same object. This adjustment is called fixation. It is sometimes described as the intersection of the visual axes at the point fixated.

(5) The oblique muscles must rotate the eyes about their axes, until corresponding retinal points occupy corresponding meridians.

The rotation adjustment is necessary because otherwise identical points of the retinæ might not correspond, even when the centres did, so that one image would appear tilted at an angle with the other.

The Horopter. Theory shows that, even when fixation is properly effected, so that corresponding retinal points occupy the same meridians, images formed on the retinæ do not necessarily fall on corresponding points. For this to be the case, it is necessary also that the objects from which those images are formed should occupy certain definite positions in relationship with one another. For example, if an object 10 feet from the eye is fixated, the images of other objects on either side will only fall on corresponding points if they lie on a circle of 5 feet radius, the centre of which is between the observer and the object fixated. For calculation shows that only then are the images formed on the two retinæ at the same distances from the centres. The form of the curve, which is called the horopter, is found to change with the different directions of the gaze. When the gaze is directed to a point on the floor, it is stated that the horopter almost corresponds with the plane of the floor.

Monocular Depth Perception. The perception of depth with the single eye is found to depend on a number of different factors, which as a rule operate together :—

(1) The apparent size of objects, the dimensions of which are known. Thus the size of a man being approximately known, his distance away is deduced from the size of the image which is formed on the retina. The further away he is, the smaller his image will appear.

(2) The colour of an object being known, the effect of distance in modifying that colour is used in depth perception. Thus trees, which when near look yellow-green, when seen at a distance through an intervening layer of haze appear blue-green or even blue. This fact is made use of by artists for expressing distance (" ærial perspective ").

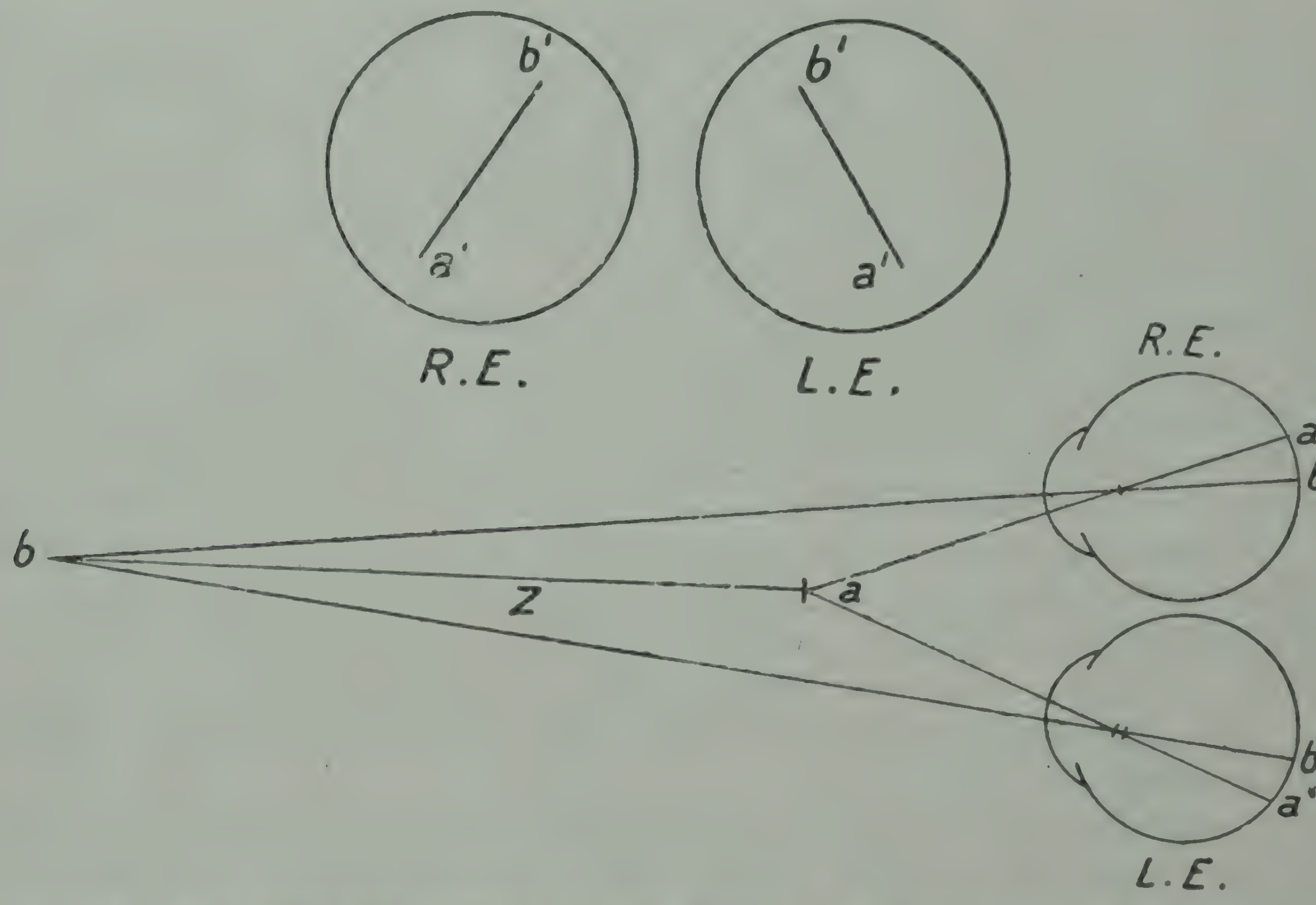


Fig. 304. The eyes are directed to the point *b*. A thread hung obliquely at *a* gives rise under these circumstances to the images shown in the upper figures —*i.e.* two images which do not lie on corresponding points. Nevertheless the thread is seen as single.

(3) The partial obstruction of a distant plane by objects nearer to the observer.

(4) The shadows which one plane casts upon another.

(5) The intensity of the light which is reflected by the object frequently varies with its shape and position. For example, the shape of a solid sphere can be accurately inferred from the distribution of intensity over its face.

(6) By linear perspective, which may be defined as the geometrical arrangement of lines in the image formed on the retina. Thus the lines of a tennis court seen diagonally from one side are all found to converge to one or other of two points on the horizon.

(7) By the intersection of objects with the horizontal plane. Thus the positions of trees in a field may be inferred with some accuracy, if the positions of the roots of the trees in relationship with the boundaries of the field be observed.

(8) By parallax, that is, the apparent movement of objects in relationship with one another. Thus if a middle plane be looked at, and the observer moves, it will be noticed that objects in a plane behind appear to move in the same direction as the observer, while those in a plane in front appear to move the opposite way. Even when we are standing still, we are all the time making involuntary movements which cause the development of parallax. This process is probably one of the most important in producing the monocular effect of depth.

(9) By the effort of accommodation required to focus an object sharply. In man the accommodation is found by experiment to give little or no perception of depth, possibly because the function is involuntary. It is thought that in birds, in which the ciliary muscles are striated and are under voluntary control, the accommodation may give valuable information of distance.

All the above factors operate together to produce an appreciation of distance which as a result of experience reaches a very high order, and with the exception of the last two, are used by the artist to produce the effect of solidity and reality. Any good picture shows us that the result can be very convincing.

STEREOSCOPIC VISION is the binocular perception of depth. It consists of all the factors which operate in the case of each eye separately, and in addition uses :—

(1) The convergence of the eye axes which is necessary in order to cause images of near objects to form on the fovea simultaneously.

(2) The dissimilarity between the images which are formed on the two retinæ (Fig. 305).

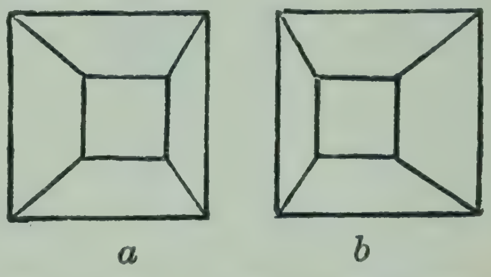


FIG. 305. To show the difference in the Images of a Truncated Pyramid as given by the right and left eyes.

That convergence has very little effect on the perception of distance can be proved by placing weak prisms, either base in or base out, in front of the eyes and in this way changing the convergence of the eye axes without changing any other condition. It is found that the apparent positions of objects are unaffected.

That there is dissimilarity between the images formed on the retina can easily be proved by experiment. Thus, if the gaze be directed towards a distant point, and the finger be held a short distance from the nose, the finger appears to be to the right of the distant point with the left eye and to the left with the right. If two photographs be taken of the same scene, but with the camera for the second photograph three inches to one side of its position for the first, it is found that, when the two negatives are placed so that objects on the horizon correspond, there is a lateral difference of position in the case of all other objects situated nearer to the camera. Measurement shows that the nearer the object the greater the difference in position. Since this is the case, it is clear that only images in one plane can be formed on corresponding retinal points ; images in all other planes must fall on points which are discrepant. Two questions therefore arise : (1) Do we see such objects doubled ? (2) If we see a single image only, is it because one of the images is displaced from consciousness by the antagonism of the other ? An answer is given by the following experiment : A Brewster's stereoscope is taken, the optical arrangement for which is shown in Fig. 306. At B and B′ two similar lantern slides are placed which show a view of any distant objects. On looking through the instrument towards the point S, the directions of the rays are changed, so that the images of the slides are seen to overlap one another. By shifting one of the slides the images may be made to fall on corresponding points of the retinæ, and they then form a single combined picture. In front of these slides is now placed another pair of slides which show the photograph of an index mark. If the indices are adjusted so that they occupy corresponding positions in relationship with the objects on the slides below them, on looking into the instrument it will be seen that these marks appear to lie in the same plane as the distant objects placed on the slides below them. If one of the index marks be moved towards the axis of the instrument, it will be seen on looking into the eyepieces that the indices now appear to lie in a plane considerably in front of their previous position, in fact that the closer they are placed together, the nearer do they appear to the observer. But the indices do not show double images unless they are moved a considerable distance together, and then the effect of distance ceases. If one of the index slides be removed and the other be moved towards and away from the axis of the instrument, the index is not found to shift its plane towards or away from the instrument. This shows that for position to be appreciated both images must be presented to consciousness simultaneously without appearing double.

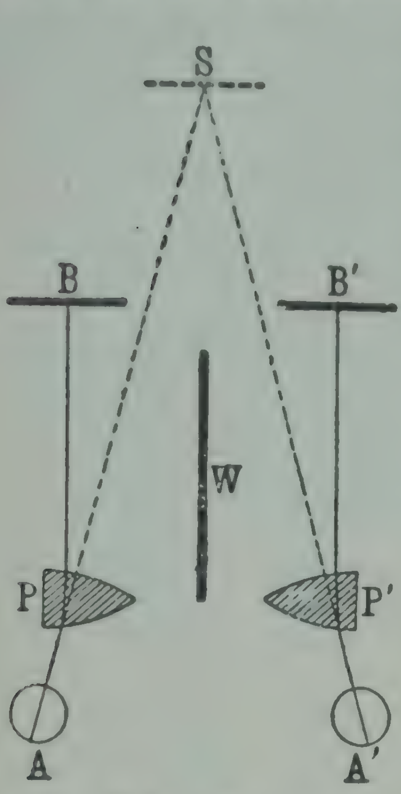


FIG. 306. Brewster's Stereoscope.

THE ACUITY OF STEREOSCOPIC VISION has been investigated in such a way that other factors which normally assist distance perception were excluded. Two methods have been used: (1) to adjust the position of a thread which lies between and parallel with two other threads until they all appear at the same distance from the observer; (2) to observe the fall of small coloured bodies of unknown size, and then to state the position of the line of fall in relationship with a fixation mark. The former method at 2 metres distance shows an average error of 1·5 mm., the latter method at the same distance an error of 40 mm. The difference between the results of the two methods is considerable; but it should be noted that in the fall method the object is seen only for 0·2 sec. If in the thread method the threads be placed horizontal it is found that the appreciation of distance is greatly impaired. The greatest acuity is found when the threads are vertical. If, however, the head is turned so that the line joining the two eyes is vertical, the greatest acuity is found when the threads are horizontal. This would be expected if the appreciation of distance is greatest when the parallax of the objects at the two eyes is greatest. Experiment shows that the recognition of position in relationship with a definite fixation mark is much more accurate than recognition of absolute distance in which there is no point of reference. Thus it is well known how inaccurate the estimation of the distance of a single source of light at night may be.

HYPOTHESIS OF DEPTH PERCEPTION. Javal's view was that the movements of the eye muscles, which are necessary in order to direct the gaze from objects in one plane to those in the next, caused impulses to travel to the brain which are interpreted in terms of distance. This view was ruled out by the fact that images which are formed on the retina for a short length of time only (0·02 sec.) can be perceived in relief.

HERING'S HYPOTHESIS was that it is the formation of similar images on points of the retinæ that do not correspond which causes distance perception. If the disparation is crossed, the object appears nearer than the fixation mark by an amount which depends on the amount of the disparation. If, on the other hand, the disparation is uncrossed, the object is recognised as being farther away. Hering supposed, moreover, that crossed disparation acts as a stimulus to convergence and accommodation, while uncrossed produces the reverse effect. We may now inquire how this hypothesis fits in with the facts. To commence with, if depth depends on disparation, it is clear that, when we perceive objects lying in different planes, we must subconsciously group them according as they fall on corresponding retinal points, or on points which are discrepant by one, two, three or more cone widths, and whether the discrepancy is crossed or uncrossed. The amount of the discrepancy must be some whole number of cone widths, because it is clearly impossible to stimulate half a cone with one impression and the other half of the same cone with a different one and obtain two distinct sensations. It is clear that space must be divided, so far as stereoscopic vision is concerned, into a number of concentric shells, the centres of which correspond with the position of the observer. Now the thickness of these shells can be readily calculated; at 1 metre they are found to be 2 mm. thick, at 10 metres 200 mm. thick, and at 100 metres 17 metres thick. If we are looking at a fixation mark 10 metres away, objects 100 mm. nearer to and 100 mm. farther from the observer will lie in the thickness of one and the same shell, and will therefore appear the same distance from the observer. Objects between 100 and 300 mm. nearer to the observer will lie in the shell corresponding to one cone discrepancy, and will therefore be appreciated as being at a different distance from the observer, appearing nearer if crossed, and farther if uncrossed. The same reasoning applies to objects at other distances. If this calculation is correct, it should be necessary to place objects more than 100 mm. from a fixation mark, which is itself placed at 10 metres, in order that a difference in the distance from the observer should be appreciated. Greeff found by experiment that $\frac{1}{50}$th the distance of the fixation mark was necessary (*i.e.* 200 mm.), the observations being instantaneous ones. If time be allowed for prolonged observation, greater accuracy in the appreciation of distance is obtainable, because different points of fixation can be used. Suppose for example that two objects 20 mm. apart be examined at a distance of 10 metres, under instantaneous observation they will appear identical as described above ; but if the examination be made more carefully, it will be found that, on fixating a point a mean distance of 100 mm. away from the objects, the distance between the two is suddenly appreciated because the demarcation between two shells now falls between them. It is in this way that the accuracy of extended observation becomes greater than that obtainable with instantaneous. The limit reached by experiment is stated to correspond to a displacement at the retina corresponding to $\frac{1}{10}$th the diameter of a cone. The corresponding values for the acuity of stereoscopic vision would be $\frac{1}{10}$th those given above, namely 0·2 mm. at 1 metre, 20 mm. at 10 metres, and 1·7 metres at 100 metres. Hering's view would therefore appear to agree well with the results of experiment. It remains to consider the type of cortical mechanism that would be necessary for the estimation of the discrepancy between the images. One type may be briefly described as follows :—To a number of parallel " planes " in the left side of the cortex are connected the terminal ends of the nerve fibres from the left halves of the two retinæ. At the middle plane fibres from exactly corresponding retinal points are connected together. At planes which lie superficially to the middle plane are connected other terminations from the same fibres but with a crossed lateral discrepancy of one cone in the 1st plane, two cones in the 2nd plane, three cones in the 3rd plane, etc. At planes which lie deep to the middle one other terminations from the same fibres are connected, but with an uncrossed lateral discrepancy of one, two, three, etc., cones as the case may be. On looking at a fixation mark on a uniform background, therefore, a series of impressions of the mark will be formed on all these planes, but in the central one only will they exactly agree, for in all the others the lateral discrepancy will cause the impressions to be duplicated. In all these other planes there will thus be antagonism, first one image and then the other predominating. When these images are combined in consciousness, the stable image from the central plane will suppress the unstable ones from all the other planes, the result being a single picture of the fixation mark. If there are in front of the fixation mark other objects lying in planes at different distances from the observer, the impulses sent up by the cones to the cortical plane will not correspond at the central plane, because their images no longer fall on corresponding points, but they will correspond in the superficial planes where the discrepancy of their images agrees with the discrepancy of the nerve-connections. These other planes will therefore predominate according as each contains the identical images, and when they are combined in consciousness these planes will suppress all the others. Since each cortical plane represents a certain lateral discrepancy, it must also represent a certain distance from the fixation mark. If consciousness recognises the plane in which a stable image is formed, it also must appreciate the distance of the object lying in that plane from the fixation mark. This would not seem any more difficult than the localisation of a touch on the skin. So far as we are able to judge there is nothing inherently impossible in the arrangement of the hypothetical cortical mechanism which has been outlined above, and therefore Hering's theory would appear to be very plausible.

CHAPTER XXVI

HEARING AND SPEECH

1. THE ORGAN OF HEARING

SOUND is conducted in the form of a wave motion. When sound is conducted through the air each sound wave consists of two phases ; during one phase the pressure in the air is above the atmospheric pressure, during the other phase it is below the atmospheric pressure. These alternate variations of pressure are conveyed to the drum skin of the middle ear *viâ* the external auditory meatus when sound waves reach an observer.

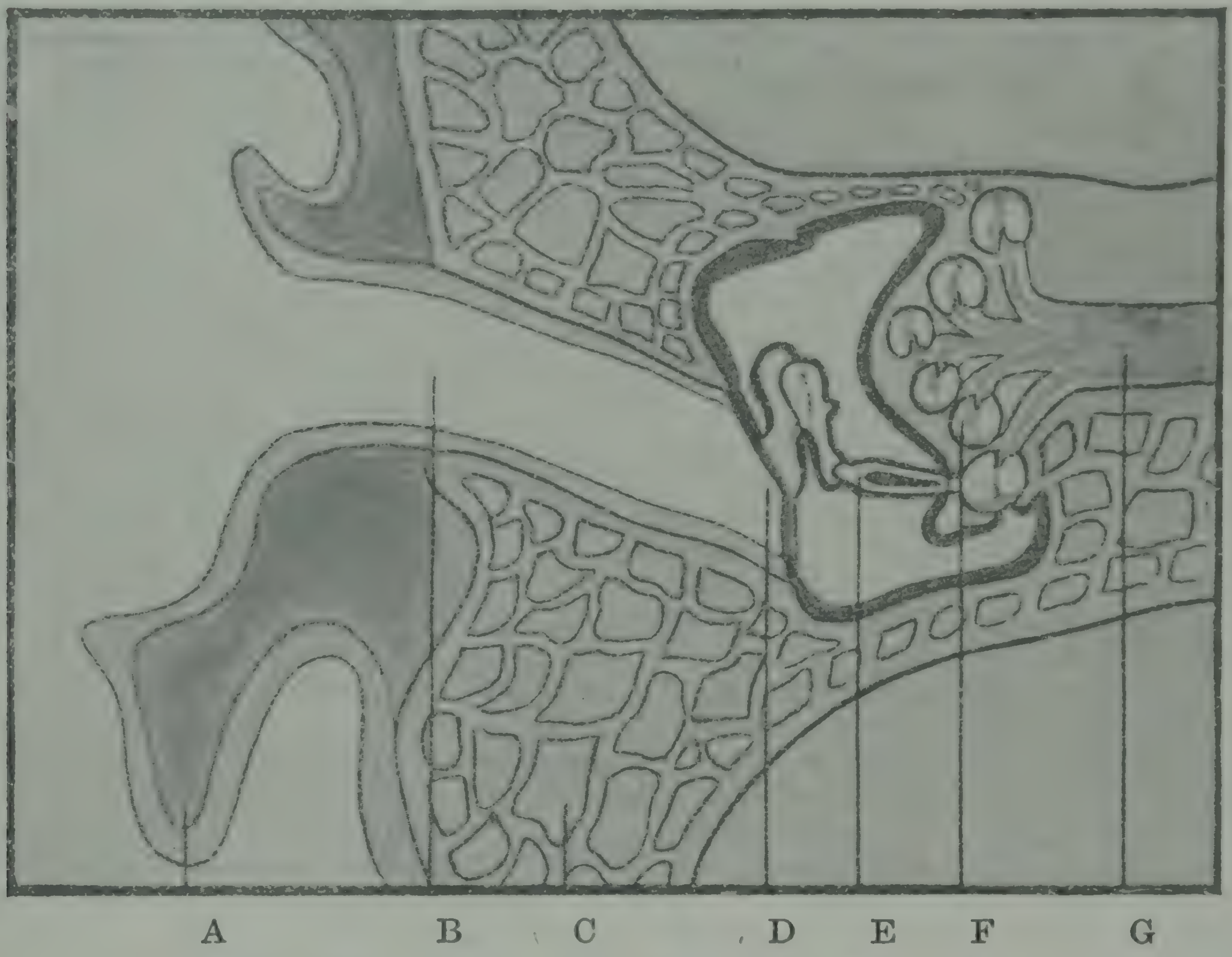


FIG. 307. Diagrammatic View of Auditory Organ.

A = Lobule of Pinna.
B = External Auditory Meatus.
C = Mastoid Air Cells.
D = Drum Skin.
E = Ossicles.
F = Cochlea.
G = Acoustic (Auditory) Nerve.

RECTIFICATION BY MEMBRANA TYMPANI. It should be noted that the drum skin has a different mode of attachment to that provided for the cone of a moving coil loud speaker. The latter is so flexibly mounted that it can move with equal facility both inwards and outwards, and so can perform a sine wave motion. The pull of the tensor tympani muscle, however, stretches the attachments of the drum skin to the meatus, so that it is only possible for the drum skin to move inwards if there is stretching of these attachments, whereas it is possible for it to move outwards by stretching the tensor tympani muscle. Of these two movements, the former is much more difficult than the latter, and thus the movements of the drum skin tend to be asymmetrical about the position of rest, so that a sine wave motion applied to it would become distorted. In consequence of this asymmetrical movement of the drum skin, which is called "rectification" by physicists, certain tones not originally conducted by the air to the ears may be generated in the ears themselves and thus be heard as if they were conducted to the ears by the air.

When a weak tone is being heard (below an intensity level of about 45 decibels) the movements of the drum skin and ossicles are linear, that is, the variations of pressure in the air cause corresponding movements according to Hooke's law. Under these conditions no ear-produced tones are heard. When a stronger tone is being heard the movements of the drum skin and ossicles become asymmetrical since Hooke's law is obeyed for an outward movement from rest, but is disobeyed for an inward movement from rest. Ear-produced tones are now heard consisting of even numbered harmonics. When a still stronger tone is being heard, such that Hooke's law is disobeyed both for movements inwards and for movements outwards, then odd numbered harmonics appear in addition to the even numbered ones.

When two pure tones are simultaneously received by an ear, then, in addition to the even harmonics for both tones and (if the tones are loud enough) the odd harmonics of those tones, it is found that there are also present both summation tones and difference tones. Suppose the two pure tones to be 700 and 1,200 cycles per second, then the following summation tones can usually be identified (700 + 1,200) = 1,900; (700 × 2) + 1,200 = 2,600; (1,200 × 2) + 700 = 3,100; (700 × 2) + (1,200 × 2) = 3,800; (700 × 3) + 1,200 = 3,300, and sometimes others as well. The following difference tones are also usually present 1,200 — 700 = 500; (700 × 2) — 1,200 = 200; (1,200 × 2) — 700 = 1,700; (1,200 × 2) — (700 × 2) = 1,000; and others generated by the ear in a similar manner. Consequently the brain receives a somewhat distorted sound picture owing to this rectification by the drum skin. This subject will be found referred to in greater detail on p. 462 under "combination tones."

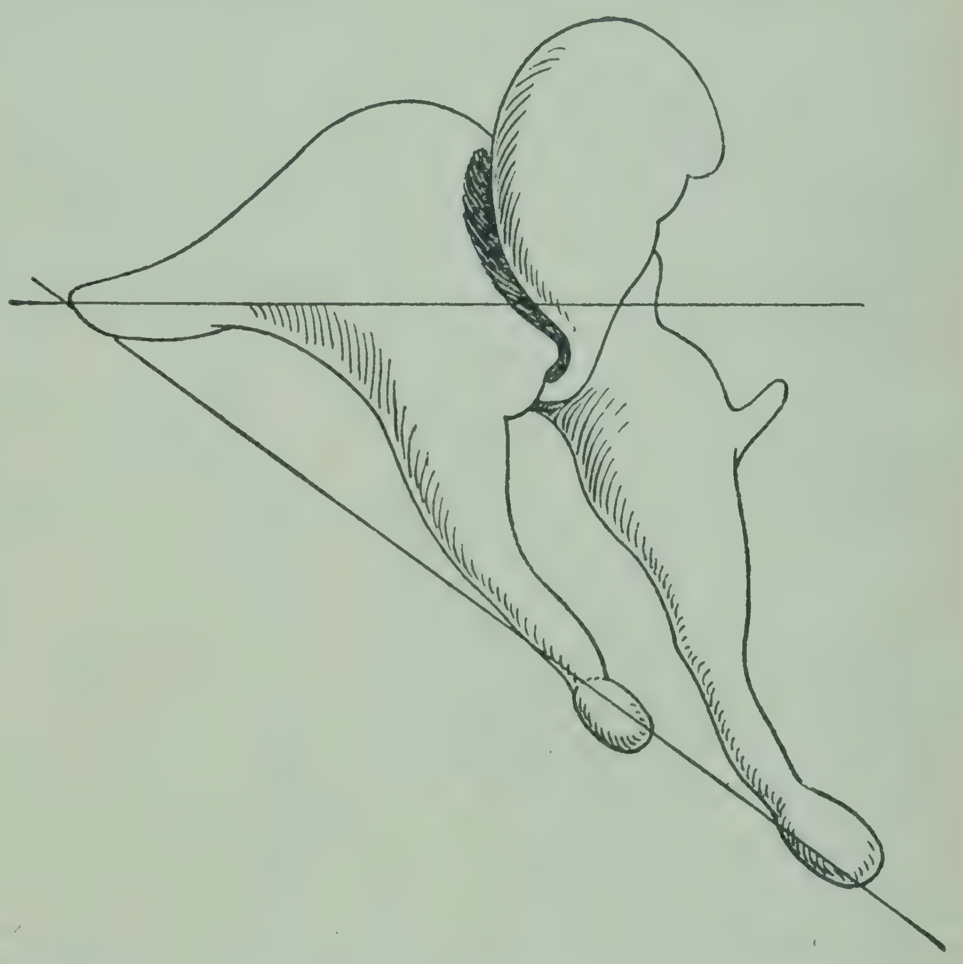


FIG. 308. To show the Relations of the Malleus and Incus to one another.

The shaded area between the two bones shows the articular surfaces which connect them. The overlapping of the two bones at the lower part of these surfaces is well shown. It is this arrangement which causes motion to be conveyed from one to the other.

THE OSSICLES. Stretching across the tympanum, from the membrana tympani to the outer wall of the internal ear, is a chain of ossicles, which are named respectively the malleus, the incus, and the stapes (Fig. 308).

THE MUSCLES found in the tympanum are the tensor tympani, which is attached to the handle of the malleus, and the stapedius attached to the base of the stapes. The tensor is innervated by a motor branch of the 5th cranial nerve, and when it is stimulated it draws the handle of the malleus inwards and so increases the tension of the tympanic membrane. At the same time the plunger of the stapes is displaced into the oval window, thus putting compression on the contents of the internal ear. The contraction of the tensor has been supposed to have a protective function and has been compared to the sphincter pupillæ (Helmholtz). Others hold that it modifies the response to low and medium tones, but even here there is a divergence of opinion because, while some hold (probably correctly) that the tensor by its contraction decreases the natural period of the drum and thus enables it to respond to rapid changes of phase and high tones, others have held the opposite view. Observation shows that contraction occurs when sounds (particularly tones of high pitch) fall on the drum, and that the contraction is bilateral, even if the stimulus be only unilateral. The reflex therefore travels *viâ* the auditory nerve to the motor centre of the 5th nerve.

Since the tensor tympani is uncontracted when no sounds are falling on the ear, it allows the drum to go slack and therefore tends to prevent this membrane from becoming stretched through being continually in tension.

The stapedius muscle is innervated by a twig from the facial nerve. Its function is problematical. Some say it antagonises the tensor by decreasing the tension on the drum; others that it reduces the tension on the contents of the internal ear by diminishing the pressure of the stapes on the oval window. Hartridge's view is that the function of this muscle is to cause the body of the incus to engage with the spur of the malleus with sufficient force to prevent chattering and lost motion when the vibrations are being transmitted from one bone to the other.

The Eustachian Canal is a tube about 35 mm. long, which connects the middle ear with the pharynx. Normally it is kept closed and hence the respiratory rhythm does not affect the pressure in the tympanum, and the noise set up by the flow of air and by the voice is not heard. When the canal is shut, the middle ear becomes a closed chamber, which appears to increase the response to low tones. Since variations in barometric pressure would not be communicated to the middle ear if the canal were always closed, the air pressure on the two sides of the drum would vary. This is avoided by a periodic opening of the canal which accompanies the acts of swallowing and yawning. When the throat is infected, the inflammation often spreads to this canal, which then becomes blocked, either by mucus or by the swelling of its mucous lining. The air in the middle ear is then gradually absorbed and the difference in air pressure on the two sides of the drum decreases its response to sound, and the affected ear thus becomes partially deaf.

Temporary deafness also occurs if the barometric pressure is suddenly altered by a rapid change of level (as in an aeroplane) or by the application of external pressure (as in a caisson). The deafness is, however, immediately relieved by swallowing, because the altered pressure is communicated to the other side of the drum through the opening of the Eustachian canal.

Functions of Tympanum. The function of the tympanic apparatus (consisting of drum, bones and muscles) is to transform the energy of the aerial vibrations incident on the drum into a series of mechanical movements of the plunger of the stapes, by which the pressure within the internal ear is rapidly varied. The evidence may be summarised as follows. (1) If in man the external ear be made to form a gas chamber which is connected with a manometric flame, the flame shows vibrations when sound falls on the drum, which could only be caused if the drum were set into vibration. (2) If the drum be gilded, and a beam of light be made to fall on it, the excursions of the beam caused by vibrations of the drum can be recorded photographically, and are found to accompany the incidence of sound waves. (3) If to the chain of ossicles a light writing lever be attached, the point of which travels over a rotating smoked drum, when sounds fall on the drum the vibrations are recorded, showing that the ossicles are set into movement. (4) By opening the middle ear from above, and sprinkling with starch grains the ossicles lying within, the movements of the different parts can be readily followed under a low power microscope. When the drum is set into vibration by sound waves, it is readily seen that the whole chain of ossicles vibrates, so as to convey the vibrations to the plunger in the oval window. Many experimenters have noted the remarkable way in which the apparatus responds to vibrations varying very greatly in rate. Tones of low or of high pitch appear to be recorded with equal impartiality and fidelity. Experiment therefore confirms our sensations, which show that the ear responds to vibrations varying from 40 to 30,000 per second. It is stated that the natural period of the ossicles and drum, owing to their small size, is very much more rapid even than $\frac{1}{40,000}$ second, and it is because of this that the system is able faithfully to respond to the vibrations of longer period concerned in audition.

Direct observation thus shows that the ossicles form levers which together conduct the vibrations from the drum to the plunger of the oval window. It is necessary to consider the effect of this lever system on the amplitude and force of the vibration. Motion is applied to the manubrium of the malleus and is communicated to the long process of the incus. The former is one and one-half times the length of the latter and therefore the stapes moves with two-thirds the amplitude of the drum. If the levers moved without friction, this would be accompanied by an increase in the force of the vibrations of one and one-half times. But owing to the air which surrounds the levers and thus damps their vibration, and to the energy required to set them in vibration on account of their mass, it is probable that the force of the vibrations which reaches the oval window is not more than half that incident on the manubrium. The drum on the other hand, has an area which is about twenty times that of the oval window and the energy incident on the drum and communicated to the manubrium is as many times greater than it would be if the sound waves were incident on the oval window direct. But owing to the energy absorbed by the levers, the magnification is probably not greater than ten times, that is one-third of the calculated amount. Two other features of the chain of ossicles should be mentioned. In the first place it will be observed that the axis, about which the malleus and incus rotate, passes through the bones, so that the big mass formed by the articular surfaces is above, and the levers below, the axis of rotation, and the ossicles are approximately balanced. Secondly, the articulation between the malleus and incus is saddle-shaped, and there is a spur on the malleus which engages with the body of the incus so that, when the tensor tympani muscle relaxes and the malleus travels outwards, the spur disengages and the incus is therefore not forced to follow. When, on the other hand, the stapedius muscle is in tonic contraction, the spur is in engagement, and vibrations are therefore communicated from one bone to the other. If, however, the force applied is excessive, owing for example to a box on the ear, then the two bones separate slightly like the limbs of a compass, and the spur passes the body of the incus without communicating the blow to it. In this way rupture of the annular seal between the plunger of the stapes and the oval window is prevented.

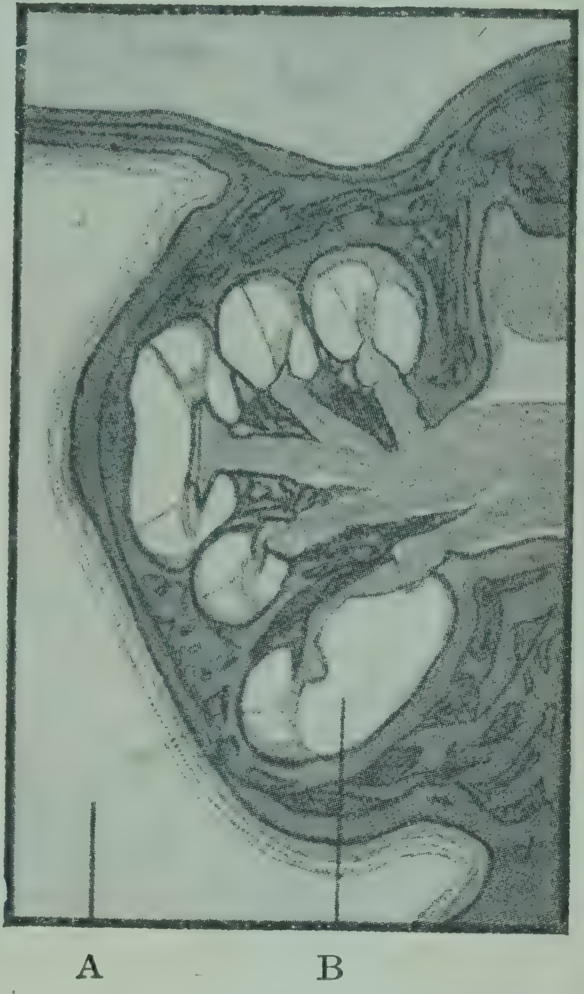


FIG. 309.
A = Middle Ear Cavity.
B = One section of Spiral duct of Cochlea.

Internal Ear. Within the petrous portion of the temporal bone are two mechanisms, anatomically in close relationship, but physiologically entirely separate. One of these mechanisms, which is called the cochlea, belongs to the auditory organ; the other, called the vestibule, consists of a series of organs which concern equilibration and have no connection with hearing.

The relation of the cochlea to the middle ear is shown in Fig. 309.

Of these structures, the cochlea alone is concerned with hearing, as the following evidence shows:

(1) Destruction of the utricle and canals causes disturbed equilibration, nystagmus and vomiting, but no deafness.

(2) Destruction of the cochlea causes deafness, but no disturbance of equilibration.

(3) Some fishes, in which no evidence of hearing can be found, possess utricle, saccule and canals, but no cochlea.

THE COCHLEA is a tube 20 to 30 mm. long, which is spirally wound round a cone of bone called the modiolus, along the axis of which runs the auditory nerve. From the modiolus a spiral lamina of bone extends about two-thirds of the way across the spiral bony canal, so as partially to divide it into two equal portions. From the outer edge of this lamina two membranes extend to the walls of the canal, so that the latter is divided throughout its length into three separate ducts. The duct nearest to the apex is called the scala vestibuli, the middle duct between the two membranes the scala media, and the duct nearest to the base the scala tympani (Fig. 309). The two membranes dividing off these ducts are quite different in structure; whereas the apical, called Reissner's membrane, is a thin layer of cells only, the basal is of complicated arrangement and is called the basilar membrane. To the latter is attached, in the organ of Corti, a series of sensitive hair cells connected with the fibres of the auditory nerve, which run through the osseous spiral lamina to the body of the modiolus. To the upper edge of the spiral lamina is attached a projecting membrane called the lamina or membrana tectoria; this is so mounted that it projects over the tips of the hair cells. Its function will be discussed later.

The way in which motion is imparted to the basilar membrane by movements of the ossicles may be described as follows. The osseous labyrinth communicates with the middle ear by means of two openings, the oval window and the circular window. The oval window connects with the upper of the three cochlear canals, *viâ* the vestibule. The upper canal is therefore called the scala vestibuli. The lower canal connects with the round window only, and since the round window is fitted with a membrane, the canal gets the name scala tympani (canal of the drum). Fitting into the oval window is the plunger of the stapes, and between the two is an annular membrane which permits motion of the plunger without allowing escape of the perilymph from the labyrinth. The circular window is also closed by its membrane, in order that leakage of perilymph may be prevented. When the plunger of the stapes is moved inwards under the influence of sound waves on the drum, the perilymph, being incompressible, is driven inwards into the vestibule. From the vestibule a corresponding volume of liquid is displaced into the scala vestibuli of the cochlea. This drives Reissner's membrane downwards, and thus increases the pressure on the endolymph in the scala media. The basilar membrane therefore moves downwards and the hair cells are drawn farther away from the membrana tectoria. But this movement of the basilar membrane presses on the fluid contents of the scala tympani which communicates with the round window,

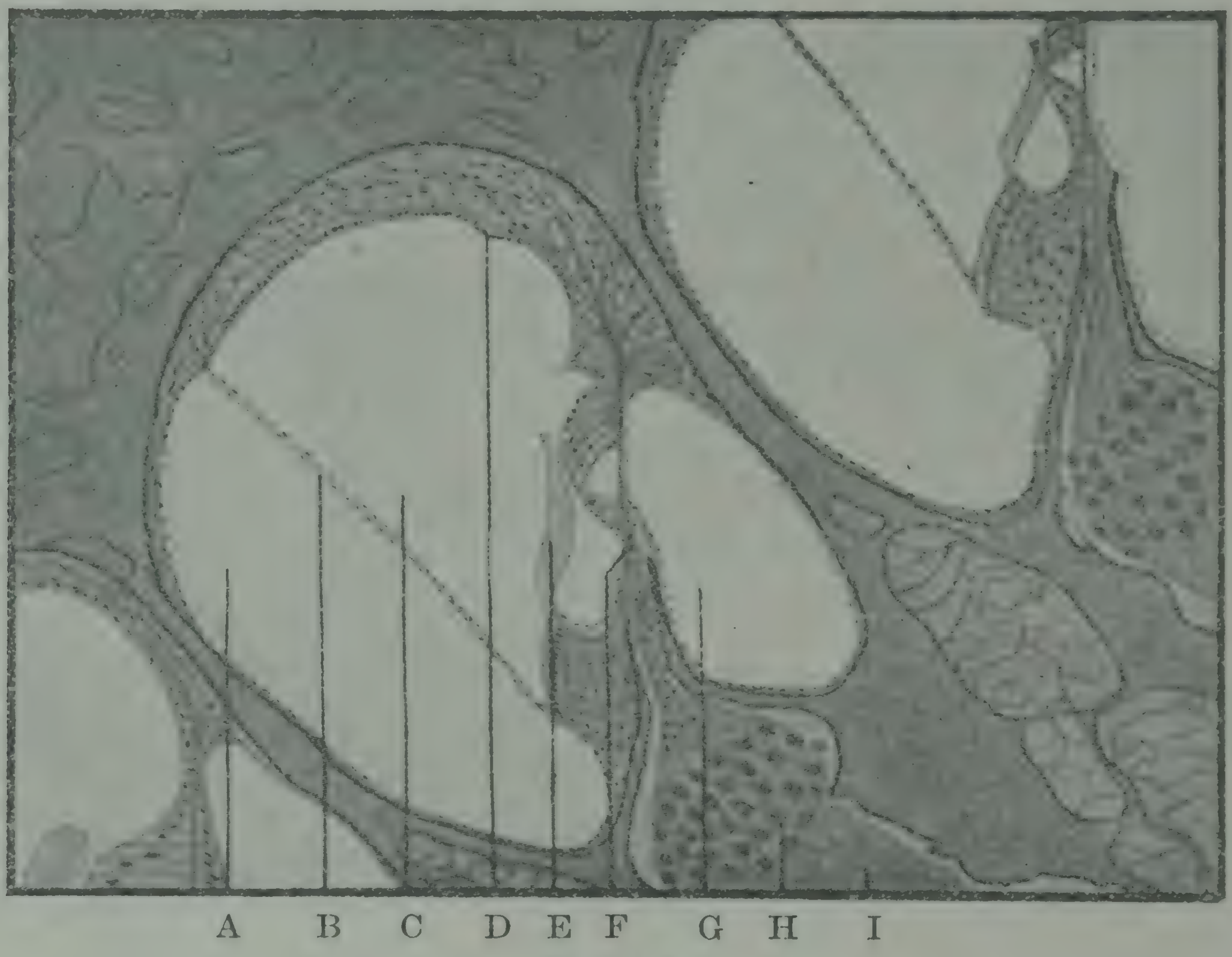


Fig. 310.

A = Scala vestibuli.
B = Reissner's Membrane.
C = Scala Media.
D = Stria Vascularis.
E = Membrana Tectoria.
F = Basilar Membrane.
G = Scala Tympani.
H = Spiral Ganglion.
I = Fibres of Acoustic (Auditory) Nerve.

and therefore causes bulging of the membrane closing that aperture.✓ Movement inwards of the plunger is therefore accompanied by movement downwards of the basilar membrane and movement outwards of the membrane of the round window and *vice versâ.* Vibrations of the ossicles are in this way communicated to the basilar membrane and to the hair cells, causing stimulation of the auditory nerve. The basilar membrane is composed partly of strong radial fibres and partly of weak longitudinal ones. It varies in width, being narrow at the base and wide at the apex.

THE ORGAN OF CORTI. The end organ of the auditory nerve is represented by the organ of Corti, which is attached to the basilar membrane (Fig. 311). It consists of a double row of stiff cells, the inner and outer rods of Corti, which run throughout the whole length of the scala media and are surrounded by sensory epithelium, the hair cells. On the inner side of the rods of Corti there is a single row, on the outer side three rows of hair cells. The fibres of the auditory nerve pass up through the column of the cochlea, through the bipolar ganglion cells which form the spiral ganglion, and then out along grooves in the spiral lamina to end in arborisations, partly in the inner hair cells and partly among the outer hair cells. Between the hair cells are the sustentacular cells, or cells of Deiters, the peripheral processes from which join together so as to form a reticulate membrane over the hair cells, the hairs themselves projecting through orifices in this membrane. Resting on the upper surface of the membrana reticularis is the membrana tectoria. To this membrane is often ascribed a damping effect on the vibrations of the structures below. Any movement of the basilar membrane would be transmitted to the rods of Corti, and by these to the overlying hair cells. With every vibration, these would oscillate in the line of their long axis, so that their hairs would move up and down and possibly strike against the under surface of the membrana tectoria.

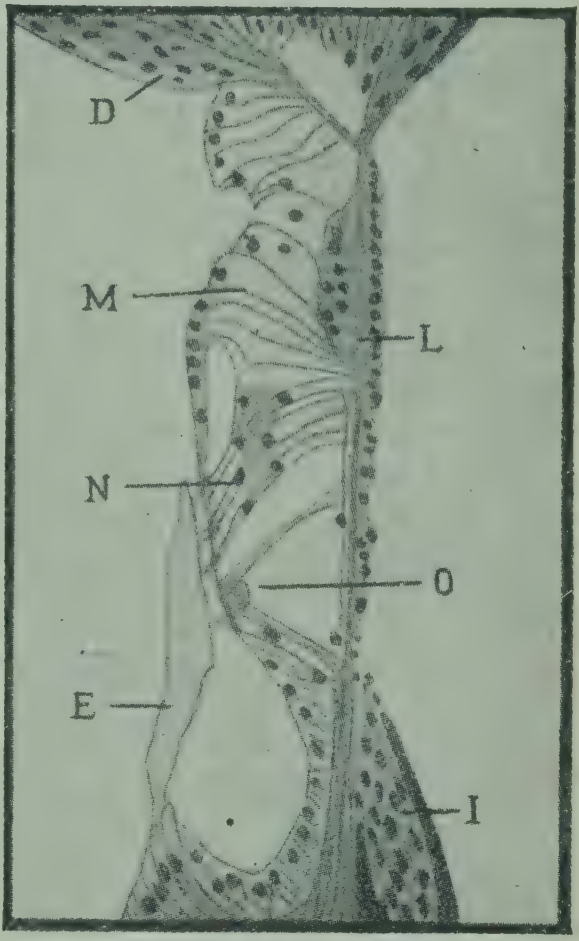


FIG. 311.

D = Stria Vascularis.
E = Membrana Tectoria.
I = Fibres of Auditory Nerve.
L = Basilar Membrane.
M = Hair Cells.
N = Hair Cells.
O = Arch of Corti.

✓Other views have, however, been advanced as to the way in which the hair cells become stimulated. Thus Wrightson states that there is a to-and-fro movement between the hair cells and the membrana tectoria, so that the hairs are bent first to one side and then to the other. Keith suggests, on the other hand, that the hairs are embedded in the membrana tectoria and that the stimuli are set up by the pulling of the hairs which must occur when the basilar membrane is moved. Further research is required to elucidate this point.

The Cochlear Microphonic Action. Wever and Bray observed that the ear acts like a microphone, that is, that when sounds fall on it variations of potential occur which are closely related to the incident sounds. Thus when a tone of 250 vibrations per second falls on the ear variations of potential occur at this rate in it. The potential variations are small and need to be amplified by thermionic valves in cascade, as for nervous tissues. They can be produced by tones as low as 5 cycles per second and by tones as high as 16,000. They start immediately the tone starts and stop when it stops. If the phase of the tone is suddenly changed, the phase of the microphonic potential also changes at the same instant (Hallpike, Hartridge and Rawden Smith). When the tones falling on the ear are loud, then harmonics can be detected in the microphonic potential curve. They are produced by distortion during transmission by the drumskin and ossicles (p. 461).

The Cochlear Telephonic Action. Andrief, Gersuni and Volokhov showed that alternating currents of audible frequencies are able to cause the perception of sounds when they are suitably applied to the ear.

In one frequently used arrangement the head is held on one side and the upper external ear is filled with warm saline, and into this is inserted a silver wire electrode. The other electrode is placed in the mouth or is strapped to the forearm. The electrodes are connected to a suitable A.C. source such as a wireless set or a beat oscillator. Tests with pure tones show that it is difficult or even impossible to obtain an intense response at any part of the musical scale because of the painful stimulation of the skin caused by the alternating electric current. Both low tones and high tones are poorly reproduced, there being an apparent cut off at about 50 cycles per second. Ear generated harmonics are strongly present, so much so that a pure tone is frequently mistaken for the octave above it. Short interruptions in a tone, or a sudden change of phase in a tone are observed as clearly as they are in an air-borne sound (Hallpike, Hartridge and Rawden Smith).

The Cause of the Electro-Acoustic Effects. It seems reasonable on physical grounds to make the assumption that the microphonic and the telephonic effects of the cochlea originate in one and the same structure. Just as in the piezo-electric effect: When a suitable crystal is moved it generates an electric potential, or when a suitable potential is applied to it a movement is brought about, so in the ear, a structure or structures are present, movements of which generate electrical potentials, or when potentials are applied from without, movements are caused which bring about the sensations of sound. What the structure or structures are is not known at the present time, the hair cells of Corti's organ have been suggested, so also has Reissner's membrane.

2. THE THEORIES OF HEARING

THE RESONANCE THEORY (of Helmholtz) states that the organ of Corti and the basilar membrane together form a series of automatically recording resonators. In the same way that each of the strings of a piano can be set into vibration by the sounding of a note which corresponds with it in pitch, so also can the different fibres of the basilar membrane vibrate to a certain note, and so cause stimulation of the hair cells which are attached to it.

Four objections have been made to this theory. (1) That the fibres of the basilar membrane are so short that they could not respond to the low notes which the ear is able to hear. The answer to this criticism is that not only the length, but also the tension and weight of a cord determine its vibration rate. In the case of the basilar membrane the tensions in the fibres are probably minute, while the weights of the arches of Corti and the hair cells must make the period of vibration so much the longer.

(2) That the separate fibres of the basilar membrane are bound together so that vibration of the separate fibres would be impossible. This objection was met by the calculation, by Helmholtz, that a uniform membrane, in which the tension was greater from side to side than longitudinally, would be able to respond in the manner required.

(3) That the difference in length of the fibres is not sufficiently great for the short ones to vibrate to notes of 4,000 vibrations per second, while the long ones vibrate to 40 vibrations per second only. This objection also fails when we reflect that not only length, but also tension and weight determine the period of vibration of a stretched cord. However accurately we can determine length and weight by histological examination, the method tells us nothing concerning tension. This objection therefore does not prevail.

(4) That if the cochlea depends for its action on the resonance of the basilar fibres, we should expect a musical note to seem to go on sounding after the note

has actually ceased. Since, on the other hand, we know from our own experience that words such as "utter," in which there is an interval of silence between the two "ts," are quite different from "udder," in which there is no interval of silence, it follows that the fibres of the basilar membrane have not been in vibration for long after the sound ceased, and therefore probably resonance of the basilar membrane is slight. But if we suppose the fibres to be highly damped, so that they come to rest at once when the note ceases, how comes it that they can be set in motion so readily that in only three or four vibrations a note is distinctly heard ? The answer is probably obtained from the fact that the cochlea is filled with liquid. This liquid makes the basilar membrane "dead beat," because movements, when the liquid is still, set up eddies which, owing to the viscosity of the liquid, quickly check the motion. On the other hand, when a sound is entering the ear, and the fluid is therefore in motion, this movement is the more rapidly imparted to the basilar membrane because of its continuity, but even if it consisted of separate fibres it would still be set promptly into vibration owing to the viscosity of the liquid. In this way one can explain both the rapid response and the rapid damping of the cochlea.

In favour of Helmholtz' theory we have the following evidence :

(1) In boiler-makers' disease we have inability to hear high notes, and we find that it is the short fibres of the basilar membrane which have degenerated.

(2) In experiments in which the ears of animals have been stimulated for long periods by the same note, subsequent examination has shown the localisation of degeneration to one part of the organ of Corti. With a high note the short fibres are affected, with a low note the long.

(3) If one of the ears be fatigued by prolonged stimulation to a constant note, its response to the same note is found to be inhibited, but notes of slightly longer or slightly shorter pitch are found to be unaffected. This shows clearly that the response to a given rate of vibration must only affect a certain limited number of hair cells and nerve fibres, and is therefore strongly in favour of Helmholtz' theory.

(4) Animals whose calls have a small range of pitch (*e.g.* birds) have short basilar membranes which vary but little in width.

(5) Animals in which different parts of the cochlea have been destroyed, appear to give definite evidence for deafness to high notes when the short basilar fibres are damaged, and deafness to low notes when the long fibres have been removed.

(6) Patients are found in whom there are islands of deafness, that is, they are deaf to a limited part of the musical scale. The Helmholtz' theory readily explains these cases as being due to local disease of certain basilar fibres or their corresponding hair cells. Further, there are cases in which the two ears give different notes, a condition called double disharmonic hearing. This is easily explained by a change in the natural period of the fibres of the basilar membrane in the diseased ear, as the result either of stretching or of increased mass due to inflammation.

(7) If the input leads of a sensitive valve amplifier be connected to the cochlea of an anæsthetised animal, then, as Wever and Bray showed, the incidence of musical tones on the ear is accompanied by potential fluctuations in the cochlea at a frequency corresponding to that of the musical tone. Hallpike and Rawden Smith found, moreover, that with low-pitched tones the potential fluctuations were greater near the apex of the cochlea, whereas, with high-pitched tones they were greater near its base, as should be the case on the resonance hypothesis.

(8) When a musical tone, transmitted by a telephone (*i.e.* an aperiodic) circuit, is suddenly changed in phase by a half-cycle the potential fluctuations in the circuit continue almost without interruption.

When a musical tone, which is causing a tuned structure to perform resonant vibrations, is suddenly changed in phase by a half-cycle the vibrations of the tuned structure are first arrested and are then re-established. On performing a similar experiment with the human ear, Hartridge found that just after the change of phase the musical tone is heard to undergo a short interruption, *i.e.* the ear, unlike a telephone, behaves as if it contained resonators. This phase-change effect was heard whether the experiment was done in an ordinary room where there are echoes or in a room with sound-absorbing walls in which echoes were reduced to a minimum ; the effect cannot therefore be due to them.

(9) When a musical tone is interrupted for short periods of time similar interruptions can be identified in a telephone circuit which is excited by the sound. When a musical tone which is causing a tuned structure to perform resonant vibrations is interrupted for a short period of time the vibrations of the tuned circuit are found to continue with only a small diminution in amplitude. On performing a similar experiment with the ear it is found that short interruptions cannot be detected, *i.e.* the ear again behaves as if it contained resonators.

(10) When two musical tones differing considerably in frequency, for example, 200 v. per sec. and 250 v. per sec., are sounded simultaneously the voltage fluctuations in a telephone circuit are found to undergo periodic increases and decreases, which correspond to beats between the two musical tones. When the above two tones are caused to set up vibrations in a graduated series of resonators, two groups of resonators are found to be set into vibration, one by each tone. A careful examination of the groups shows that neither of them is exhibiting appreciable beats. On performing the experiments with the ear no beats are to be observed, *i.e.* the ear behaves as a series of resonators.

(11) When two musical tones having frequencies in the ratio of 1 to 3 are sounded simultaneously the voltage fluctuations in a telephone circuit excited by them are found to be greatly affected by the relative phases between the two tones. In one phase relationship a single sharp peak curve is produced. In another phase relationship a double hump curve is produced. When the above two tones are caused to set up vibrations in a graduated series of resonators two groups of resonators are found to be set into vibration, one group for each tone. The careful examination of the groups shows that neither their frequency nor amplitude is affected appreciably by changing the relative phases of the two tones. On performing the experiment with the ear it is found that changing the relative phases of the tones has no effect on the sounds which are produced, *i.e.* the ear behaves in accordance with the resonance theory and contrary to the telephone theory.

(12) It is well known that resonators exhibit the properties of selectivity (sharpness of resonance) and persistence, and that both these properties are related to the frictional losses to which the resonators are subject. In consequence of this, when different degrees of selectivity are multiplied by corresponding degrees of persistence a constant value is obtained, as is shown in the following table :—

S (selectivity)	P (persistence)	S × P
1·5	40	60
2·0	30	60
3·0	20	60
4·0	15	60
6·0	10	60

(Selectivity in the above table = percentage difference of frequency causing 1/10th the amplitude of that caused by the intune frequency at the same energy level.)

(Persistence in the above table = number of vibrations which correspond to the time required for the amplitude to be reduced to 1/10th its initial value.)

Now the selectivity of a graduated series of resonators can be ascertained experimentally by finding the smallest difference of frequency which causes beats owing to the overlap of the two vibrating groups of resonators. The persistence of a series of resonators can be ascertained experimentally by finding the duration of the longest interruption in the incoming sound energy which produces no observable decrease in amplitude. Both these methods of research are applicable to the ear, and the table below gives the values I obtained. The last column shows that their product gives a reasonably constant value, as should be the case if the ear resonators are of the kind familiar to the physicist.

Tone	S	P	S × P
128	12·7	1·78	22·6
256	10·0	2·06	20·6
326	9·45	2·19	20·7
384	9·07	2·18	19·75
512	8·45	2·37	20·0
640	8·15	2·54	20·65
760	7·82	2·68	20·95
1024	7·22	3·01	21·70

It would appear, therefore, that the evidence in favour of Helmholtz' theory is very convincing. Other hypotheses have been proposed, however, two of which will now receive brief consideration.

The Telephone Theory (of Rutherford) compares the cochlea with a telephone. In the same way as the diaphragm of the receiver is set into vibration by the sound waves, and starts corresponding variations in the strength of the current conducted to the transmitter, so the vibration of the basilar membrane as a whole causes impulses to be sent up the auditory nerve which correspond with the air vibrations received by the ear. Analysis does not take place in the cochlea at all, but in the brain. Wrightson, who has restated this theory and added much detail to it, states that the cerebral analysis is effected by differences between the time intervals of the points of zero pressure and of the maximum plus and minus pressures.

The following objections may be stated against the Rutherford-Wrightson hypothesis :

(1) It assumes that the auditory nerve can conduct complicated wave forms, intact as to pitch and amplitude, at rates up to 15,000 vibrations per second. Rutherford, in this connection, pointed to the motor nerves of the bee's wing, which are capable of responding to transmitted impulses at 460 per second. Between 15,000 and 460 is, however, a big gap which will certainly have to be bridged before this view as to the transmission of the vibrations intact to the brain can be accepted.

(2) We cannot picture a cerebral apparatus which can analyse these complicated nerve impulses even if they could reach it, and neither Rutherford nor Wrightson assist us to do so. The relegation of the powers of analysis to the cerebral cortex is, at the present at any rate, equivalent to giving up any attempt to explain the power of analysis possessed by the organ of hearing.

(3) It would seem that a very much simpler organ than that of the cochlea would be sufficient to convert sound waves into nerve impulses, if no analysis of the stimulus took place there.

(4) It would be very difficult to explain, on this hypothesis, the localisation of deafness to certain notes, which accompanies disease of part of the organ of Corti.

(5) This hypothesis does not explain why fatigue to one note leaves the response to all other notes apparently unaffected in intensity.

The objections to this hypothesis are therefore of a formidable character. Much additional evidence in its favour would be necessary to place it even on a par with the theory of Helmholtz.

The Sound Pattern Theory (of Waller and Ewald), stated that the basilar membrane vibrated in the form of pressure patterns, which are similar to those which may be seen on a vibrating plate. Ewald found by experiment that the patterns take the form of equidistant stationary nodes or ridges, the distance between which varies with the pitch of the note entering the mechanism. The distance between the nodes is measured by the hair cells, and corresponding impulses are sent to the auditory centre. The advantage of this hypothesis is that like that of Helmholtz, it places the analysis of the sound waves in the cochlea and therefore does not, like Rutherford's hypothesis, require the transmission by the auditory nerve of rapid impulses or the analysis of such impulses by the brain. It is clear, however, that so far as our present knowledge goes the evidence is all in favour of the Helmholtz view.

The Volley Theory has been devised to explain two different facts : (1) that the position in space of sources emitting tones below about 1,000 cycles is identified by the recognition of the time interval between the arrival of the sound at the two ears. When the source is straight in front of the observer the sound reaches the right and left ear simultaneously. If this source is moved round to the right the sound reaches the right ear first. If the brain is to appreciate this fact the nerve impulses reaching it must be discontinuous and must consist of volleys having time relationships corresponding to the vibrations set up in the ear by the sound waves. Thus a tone of 500 vibrations per second from a source of sound 45° right of the observer will cause 500 volleys of nerve impulses to travel to the brain along each auditory nerve. But the sound waves to the left ear have to travel about 4 inches farther than those to the right, and the velocity of sound in air being 13,200 inches per second, this distance will take $\frac{1}{3,300}$ of a second and the volleys up the right auditory nerve will be that time ahead of those up the left nerve. (2) That the responses in the auditory nerve are found to correspond in frequency to those of the tones falling on the ear up to about 2,200 cycles per second. But the separate fibres of the auditory nerve, like those of other nerves which have been experimentally investigated, cannot transmit more than about 1,000 impulses per second. The volley theory explains the faster rate of transmission by supposing that each fibre responds to every other sound wave, and possibly when the fastest rate of all is being transmitted (2,200 cycles per second), each fibre of the auditory nerve may respond to every third sound wave. When, however, tones of higher pitch are falling on the ear and each of the auditory nerve fibres should be transmitting every fourth, fifth, sixth, etc., up to possibly every tenth wave, the transmission of the nerve impulses has become too irregular for the precise conditions to be experimentally identified. If all the auditory nerve fibres had precisely the same refractory phase, and the same rate of conduction, then proof of the volley theory could probably be obtained for all musical tones including the highest. But this seems not to be the case. That the volley theory holds good for tones below about 2,400 cycles per second was shown by the experiments of Derbyshire and Davis. In Fig. 312 is shown the relationship between the frequency of the tone and the size of the electric response of the auditory nerve of a cat taken immediately after the brief transcent (the " on effect ") which always occurs at the beginning of a musical tone. It will be seen that the response is a maximum from 500 to 1,000 cycles per second. Experiment shows that this maximum extends right down the musical scale to low tones. Above 1,000 cycles per second, the graph shows a fall so that at 1,400 cycles per second the response has half the maximum value. At about 1,600 cycles the nerve response has flattened out and at 1,800 undergoes a second abrupt fall. The first fall at 1,100 cycles occurs at a point where, on the volley theory, a fall would be expected because some of the auditory nerve fibres are beginning to respond to every other wave. The second fall at about 1,800 cycles occurs at a point where on the volley theory a further fall would be expected, because some of the auditory nerve fibres are failing to be able to respond to every other wave and are beginning to respond to every third wave.

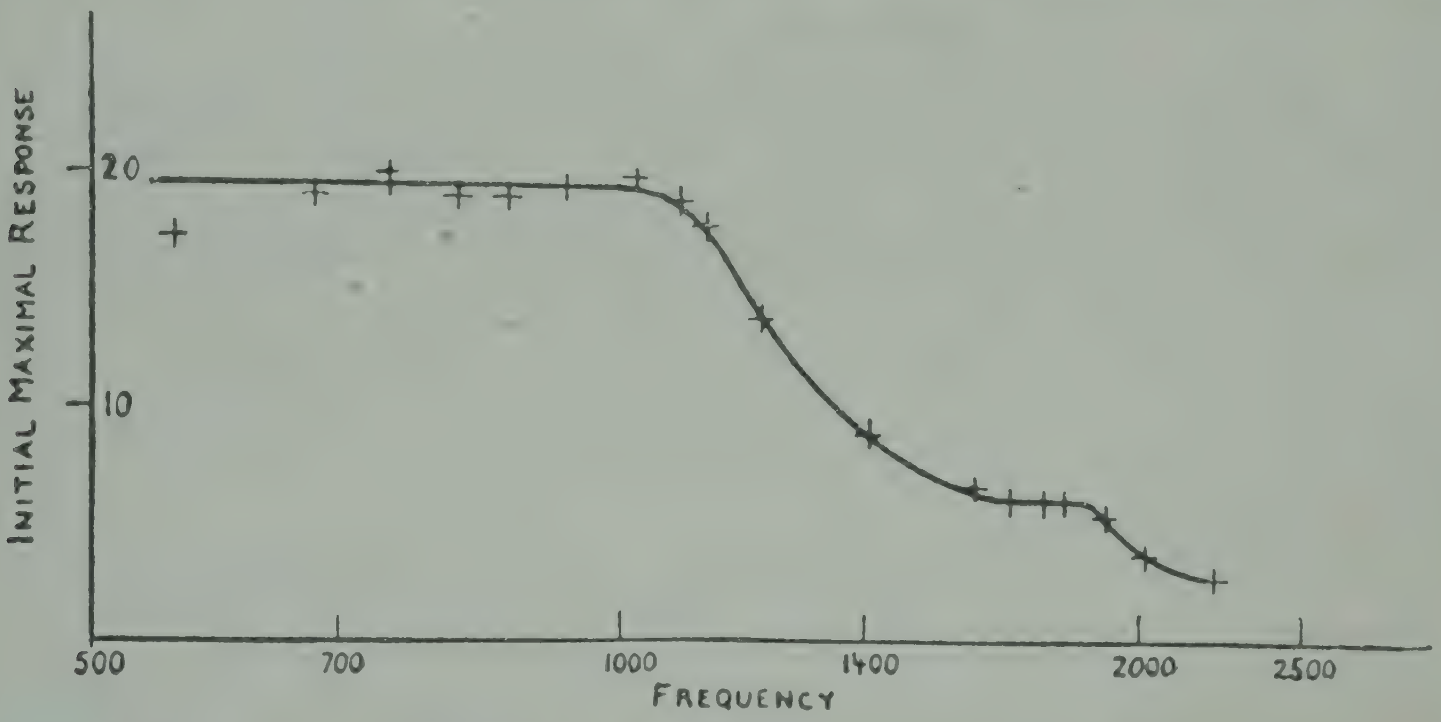


FIG. 312. Shows the relationship between the electrical response of the auditory nerve and frequency in cycles per second. (Derbyshire and Davis, *Amer. J. Phys.*) Between 500 and 1,000 cycles the curve is level because all the auditory nerve fibres respond to every vibration of the basilar membrane. The bend in the curve that occurs at 1,000 cycles is due to the fact that above this frequency, half the nerve fibres are responding to every other vibration in accordance with the volley theory. The bend in the curve that occurs at a frequency of 1,900 cycles is due to the fact that above this frequency, one-third of the nerve fibres are responding to every third vibration.

3. THE PHENOMENA OF HEARING

Limits of Pitch. The ear is unable to perceive a tone the pitch of which falls above or below certain fairly well-defined limits. If the number of vibrations is less than about twenty per second no musical tone is produced, the individual vibrations being perceived as a series of pulses in the surrounding air, and it is only when we increase the number to about thirty per second that we are able to appreciate a pitch in the note produced. As the number of vibrations per second is increased the note rises steadily without break till we arrive at about 15,000 vibrations per second. Above this number of vibrations, the human ear is incapable of perceiving any note at all, though it is probable that small animals can hear notes still higher in the scale. In music neither the lowest nor the highest tones are used. The lowest tone of large organs, 16 vibrations per second, is that produced by the sixty-four foot pipe, and one can hardly speak of its effect as that of a musical tone. The highest notes employed in music are *a*4 and *c*5 with 3,520 and 4,224 vibrations per second on the piano, and *d*5 with 4,752 vibrations on the piccolo flute. In music therefore we only employ between 25 and 4,800 vibrations per second, *i.e.* about seven and a half octaves.

Beats and Dissonance. If we sound two tuning forks, the vibrations of which differ only by one or two per second, the phenomenon known as " beats " is produced. This is due to the summation and interference of the waves from the two tuning forks. Let us suppose we have tuning forks vibrating, one at 100 and the other at 101 times a second, and that they begin vibrating together. At first the waves of compression started by each fork will coincide, so that the total compression of the air at each beat will be the compound effect of the compression produced by the two forks, which will reinforce one another. After the lapse of half a second the tuning forks will be at different phases of their excursion. The 101 fork will be moving in one direction while the 100 fork is moving in the other, so that the compression produced by one fork coincides with the expansion of the air produced by the moving backwards of the other fork. The sound produced by one fork is therefore diminished by the sound produced by the other fork, and the total sound is less than that of either of the two forks. At the end of one second, the phases of the two forks once more corresponding, we shall get the sound increased in loudness ; thus there is an alternate waxing and waning of the sound which recurs once a second and is spoken of as a " beat."

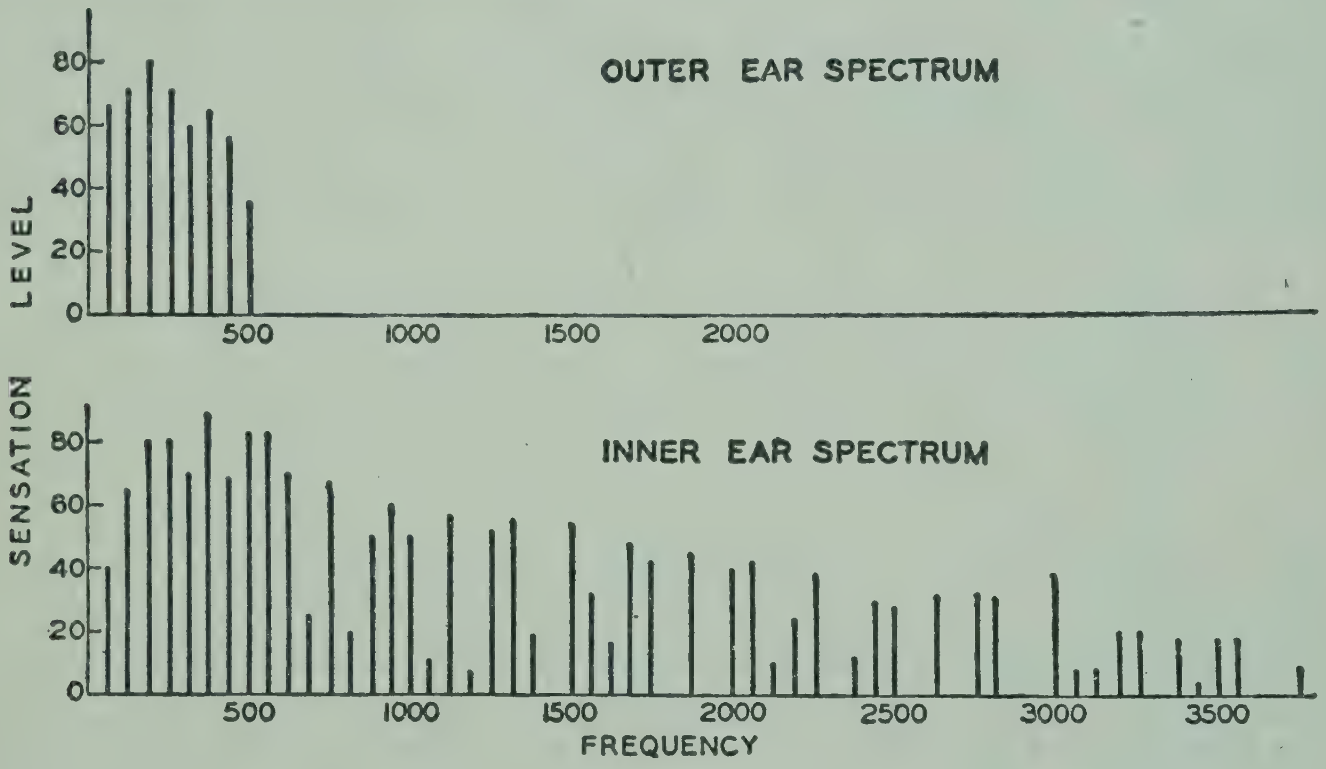


Fig. 313. Loudness spectrum : The top graph shows actual spectrum of a note of an organ pipe impressed upon the outer ear. The lower one gives estimated spectrum as impressed upon the nerve endings of the cochlea. The additional frequencies introduced owing to rectification effected by the drum are well known. (From Harvey Fletcher, " Speech and Hearing," Macmillan & Co., London, 1929.)

The number of beats per second may be used to determine the differences in the vibration frequencies of two forks. Thus, two forks vibrating, one at 100 and the other at 110, will give ten beats per second. As the number of beats increases, the effect produced on the ear becomes more and more disagreeable, just as the rapid alternation of illumination produced by a flickering light is disagreeable to the eye. This objectionable character of the sound is most marked when the beats recur at about thirty-three times per second; the individual beats are not then distinguished, but we speak of the sound as discordant or *dissonant.*

COMBINATION TONES. If two tuning forks, with an interval of one-fifth between them, are sounded together, we may hear a weak lower tone, the pitch of which is an octave below that of the lower fork. This is known as a "combination tone." The combination tones are divided into two classes: (1) "difference tones," in which the frequency is the difference of the frequencies of the generating tones; (2) "summation tones," which have a pitch corresponding to the sum of the vibrations of the tone of which they are composed. By means of appropriate resonators these tones can be reinforced, showing that they have an objective existence. They are, however, reinforced in the ear by the asymmetry of the drum skin and ossicles.

In consequence of these combination tones the sound spectrum of the motion conveyed to the cochlea is much richer in the higher frequencies than is the wave motion incident on the drum.

CONSONANCE. The opposite condition of *consonance* or harmony involves therefore, in the first place, an absence of beats, *i.e.* of rhythmic oscillations of amplitude of sound waves which reach the ear. The most complete consonance is attained when the two notes which are sounded together are identical in frequency. Almost as complete is the consonance obtained when a note is sounded together with its octave. The octave and other consonant intervals which are employed in music are as follows:—

1:2	Octave
2:3	Fifth
3:4	Fourth
4:5	Major third
5:6	Minor third
5:8	Minor sixth
3:5	Major sixth

It will be noticed that in all these consonant combinations, the vibration frequencies of the notes are in proportion to small whole numbers. If we put down not only the fundamental tones of these notes, but also their overtones, we shall see that there is considerable identity as regards the latter. In the case of the octave the two are almost identical, the only difference being the ground tone of the lower note; but the identity diminishes as we pass from the octave through the thirds to the sixths. The overtones which are identical are shown by black type:

Overtones

$\frac{1}{2}$		**1**	.	**2**	.	3	.	**4**	.	5	.	**6**	.	7	.	**8**	.	9	.	**10**	
		D		**2**		S		**4**				**6**				**8**				**10**	
$\frac{2}{3}$		2	.	4	.	**6**	.	8	.	10	.	**12**	.	14	.	16	.	**18**	.	20	
	D		3		S	**6**			9			**12**			15			**18**			
$\frac{3}{4}$		3	.	6	.	9	.	**12**	.	15	.	18	.	21	.	**24**	.	27	.	30	
	D		4		S 8			**12**			16		20			**24**			28		
$\frac{4}{5}$		4	.	8	.	12	.	16	.	**20**	.	24	.	28	.	32	.	36	.	**40**	
	D		5		S 10		15			**20**			25		30		35			**40**	
$\frac{5}{6}$		5	.	10	.	15	.	20	.	25	.	**30**	.	35	.	40	.	45	.	50	
	D		6		S 12		18		24			**30**			36		42		48		
$\frac{3}{5}$		3	.	6	.	9	.	12	.	**15**	.	18	.	21	.	24	.	27	.	**30**	
	D		5		S		10			**15**			20				25			**30**	
$\frac{8}{9}$		8	.	16	.	24	.	32	.	40	.	48	.	56	.	64	.	**72**	.	80	
	D		9		S 18		27		36		45		54		63			**72**			
$\frac{8}{15}$		8	.	16	.	24	.	32	.	40	.	48	.	56	.	64	.	72	.	80	
	D	15			S		30				45				60				75		90

D = difference tone. S = summation tone.

It will be observed that in all cases, with the exception of the octave, the summation tone is discordant with the other partial tones.

In the second $\left(\frac{8}{9}\right)$ and the seventh, both of which are discordant, it is only the eighth and ninth overtones which are identical, while the fundamental tones will as a rule be so close together that beats of a number calculated to give dissonance will be produced. Since the phenomenon of beats depends on the absolute number of vibrations per second, they are more easily produced by two notes near together at the lower end of the scale than at the upper end. Thus the dissonance is quite perceptible in a major third at the lower end of the piano, but disappears at the upper part, since here the beats produced are so rapid that they become imperceptible.

The various notes used in music are obtained by employing the consonant intervals which we have given above. The major chord is composed of the fundamental tone, the major third and the fifth. If we take "*c*" as the fundamental tone, the notes of the chord are *c*, *e*, *g*, with vibration frequencies corresponding to 1, $\frac{4}{5}$, $\frac{3}{2}$, *i.e.* 4, 5, 6. The major chord from *g* is *g*, *b*, *d*, *i.e.* three notes with vibration frequencies corresponding to $\frac{3}{2}$, $\frac{15}{8}$, $\frac{9}{4}$, *i.e.* 4, 5, 6. The major chord from the fourth, *f*, is *f*, *a*, *c*, with the vibration frequencies $\frac{4}{3}$, $\frac{20}{12}$, $\frac{12}{6}$, *i.e.* 4, 5, 6. The C major scale is therefore as follows :

C	D	E	F	G	A	B	C	
1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	2	Major scale
1	$\frac{9}{8}$	$\frac{6}{5}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{8}{5}$	$\frac{15}{8}$	2	Minor scale

Different instruments are tuned to one normal note, *i.e.* to A with 440 vibrations per second (this note varies somewhat in different countries). Taking this as the normal, the vibration frequencies of the various notes used in music are given in the following Table :—

Notes.	Vibrations per second.						
C. . .	33	66	132	264	528	1056	2112
D. . .	37·125	74·25	148·5	297	594	1188	2376
E. . .	41·25	82·5	165	330	660	1320	2640
F. . .	44	88	176	352	704	1408	2816
G. . .	49·5	99	198	396	792	1584	3168
A. . .	55	110	220	440	880	1760	3520
B. . .	61·875	123·75	247·5	495	990	1980	3960

Sound Analysis. Not only can the ear appreciate differences between different musical instruments, dependent on the varying overtones present in the sound produced by each instrument but, when a number of these instruments are sounded simultaneously, the ear can pick out from the compound sound the notes due to an individual instrument ; and a person with a trained ear can with ease name notes composing any chord struck on an instrument such as the piano.

This power of analysis, which is possessed by the auditory apparatus, may be stated in the form of a law, known as Ohm's law :

" Every motion of the air which corresponds to a composite mass of musical tones is capable of being analysed into a sum of simple pendular vibrations, and to each single vibration corresponds a simple tone, sensible to the ear and having a pitch determined by the periodic time of the corresponding motion of the air."

Sound Localisation. Experiment shows that man and animals can appreciate with fair accuracy the direction from which a sound is coming. There has been considerable speculation as to how this information is obtained, and, although the subject has not been completely elucidated, it appears to have been established that the following factors are important :—

(1) The intensity of the sounds entering the two ears. When a sound is coming from one side, the ear on that side receives the more powerful stimulus.

The relative intensities of the components of high and low pitch vary with direction, because the notes of long wavelength (low pitch) will be diffracted the more readily round the head to the ear away from the sound than will those of short wavelength (high pitch).

(2) The sounds reaching the nearer ear will arrive earlier than those stimulating the other, because of the time taken to travel round the head. The nerve impulses from the two ears will therefore not arrive at the same instant, and by an appreciation of the difference in time, the approximate position of an external object can be gauged.

That this factor is of great importance can be shown by experiment in the following manner : A stethoscope with two earpieces is fitted in position, and to its mouthpiece is applied a loud tuning fork. The tube connecting the mouthpiece to one of the earpieces has an adjustable U-piece like a trombone, so that the distance travelled by the sound in reaching that ear can be varied. The other tube has a length which is equal to that of the others, with the U-piece in its mid position. When a note is sounded and the U-piece altered, the position of the sound appears to move from one side to the other according to which ear has the shorter tube.

In everyday life both the intensity and the time factors play their part. The intensity factor is more important for frequencies above 1,000 v. per sec., the time factor being more important for frequencies below this. Thus in the case of a musical tone of 400 v. per sec. rich in harmonics, the fundamental (400) and first harmonic (800) would involve the time factor while the higher harmonics would be bringing in the intensity factor.

(3) In animals, the ability to turn the ears in different directions and so find the direction of maximum intensity must be of the utmost possible value in sound localisation.

Note on the Decibel. The DECIBEL is a physical unit in terms of which DIFFERENCES IN SOUND INTENSITIES can be expressed. 10 DECIBELS = 1 BEL, called after Alexander Graham Bell, who invented the telephone. The decibel scale is logarithmic. INTENSITY DIFFERENCE BETWEEN TWO SOUNDS IN DECIBELS =

$$10 \times \log. \frac{\text{physical power of one sound}}{\text{physical power of other sound}}$$

e.g., sound 1 : 2 = 3 decibels.
1 : 5 = 7 ,,
1 : 100 = 20 ,,
1 : 1,000 = 30 ,,

A scale of loudness has been constructed as follows :—

The faintest audible sound	0
Rustle of leaves in gentle breeze	1 bel.
Whisper, 4 feet	2 bels.
Quiet suburban street, London	3 ,,
Centre of New York at quietest time of night	4 ,,
Conversational voice, 12 feet	5 ,,
Busy traffic, London	6 ,,
Loud peal of thunder	7 ,,
Very heavy street traffic, New York	8 ,,
Lion roaring, 18 feet	9 ,,
Riveter, 35 feet	10 ,,
Steel plate hammered by four men	11 ,,
	12 ,,
Limit of ear's endurance	13 ,,

For Reference

BEATTY (1932). "The Hearing of Men and Animals" (London : George Bell.)
HARVEY FLETCHER (1929). "Speech and Hearing." (London : Macmillan & Co.)

KERRIDGE (1938). Recent Advances in Knowledge concerning Hearing and Speech. *Physiol. Rev.*, **18**, 59.

RANKE (1935). Physiologie der Schnecke und des Cortischen Organs. *Ezebn. d. Physiol.* **37**, 12.

STEVENS and DAVIS (1938). "Hearing." (London: Chapman and Hall.) ✓

4. SPEECH

The development of the analytical powers of the auditory apparatus is closely connected with that of the faculty of speech, which is best considered with the central nervous system. We may first consider the mechanism of production of *voice*, which man shares with many other animals, before discussing the physiology of the wholly human faculty of speech.

Voice is produced in the larynx, a modified portion of the wind-pipe, by the vibrations of two elastic bands, the vocal cords, which are set into action by an expiratory current of air from the lungs. In many respects the larynx resembles a reed instrument, in which a current of air is caused to vibrate by the vibrations of an elastic tongue. Whereas, however, the period of the vibrations in such an instrument, and therefore the note, is determined by the length of the tube which is attached to the reed and by the lengths of the reeds themselves, in the larynx the note produced by the blast of air is modified partly by alterations in the tension of the vocal cords, partly by altering the size of the aperture and partly by varying the strength of the blast of air.

ANATOMICAL MECHANISM OF THE LARYNX. The essential framework of the larynx is formed by four cartilages, viz. the cricoid, the thyroid, and the two arytenoid cartilages (Fig. 314). The cricoid cartilage, which lies immediately over the uppermost ring of the trachea, is shaped like a signet ring, the small narrow part being directed forwards and the broad plate backwards. The thyroid cartilage consists of two parts or alæ, joined together in front and forming the prominence known as Adam's apple; behind, it presents four processes or cornua, the superior of which are attached by ligaments to the hyoid bone, while the inferior cornua articulate with the postero-lateral portion of the cricoid cartilage. By means of this articulation very free movement is permitted between the two cartilages, the general direction of movement being one of rotation of the cricoid cartilage on the thyroid, round a horizontal axis directly through the two articular surfaces between the two cartilages, while movements of the thyroid upon the cricoid are also possible in the upward, downward, forward, and backward directions. The two arytenoid cartilages are pyramidal in shape. By their bases they articulate at some distance from the middle line with convex articular surfaces situated in the upper margin of the plate of the cricoid cartilage. The anterior angle of the base is the vocal process, while the external angle is the muscular process of the arytenoid. The crico-arytenoid joints permit of two kinds of movements of the arytenoid cartilages:

(1) Rotation on their base around their vertical long axis, so that the anterior process is rotated outwards and the muscular process backwards and inwards, or conversely.

(2) Sliding movements of the whole arytenoid cartilage either outwards or inwards, so that their inner margins may be drawn apart or approximated.

The larynx is covered internally by a mucous membrane continuous with that of the trachea. It is lined with ciliated epithelium, except over the vocal cords, where the epithelium is stratified. The two vocal cords, or thyro-arytenoid ligaments, consist of elastic fibres which run from the middle of the inner angle of the thyroid cartilage to be inserted into the anterior angle of the arytenoid cartilages. Their length in man is about 15 mm., in woman about 11 mm. The cleft between them is known as the glottis or *rim glottidis*.

Two ridges of mucous membrane above and parallel to the vocal cords are the false vocal cords. Between the true and the false vocal cords on each side is a recess known as the ventricle of Morgagni. This ventricle permits the free vibration of the vocal cords. The false cords take no part in phonation, but help to keep the true cords moistened by secretion of the numerous mucous glands with which they are provided. The false cords are also used in holding the breath. For this purpose they function in a similar manner to the mitral valve of the heart. It is found that animals, who need the thorax to be fixed in order that they may climb or strike, have well-developed false cords. The position and tension of the vocal cords are determined by the action of the intrinsic muscles of the larynx. The part taken by the various muscles in each movement cannot be directly ascertained. We can in most cases study only the direction of the fibres and judge, from this direction and consequent isolated action of the muscles, the part taken by any given muscle in the production of voice. The chief muscles (Fig. 314) are the following :—

(1) The *crico-thyroid* muscle is a short triangular muscle attached below to the cricoid cartilage and above to the inferior border of the thyroid cartilage: the fibres pass from below upwards and backwards. When this muscle contracts, the cricoid cartilage is drawn up under the anterior part of the thyroid cartilage, so that its broad expansion behind, with the arytenoid cartilages, is drawn downwards and backwards, thus putting the vocal cords on the stretch. This muscle is probably the most important in determining the tension of the vocal cord.

(2) The *posterior crico-arytenoid* muscle arises from a broad depression on the corresponding half of the posterior surface of the cricoid cartilage. It passes upwards and outwards, its fibres converging, to be inserted into the outer angle of the arytenoid cartilage. These muscles rotate the outer angle of the arytenoid cartilages backwards and inwards. They thus cause a movement outwards of the anterior angles, so that the glottis is widened. During every act of inspiration there is a widening of the glottis, which is probably effected by contraction of these muscles. If they are paralysed the vocal cords are approximated and tend to come together during inspiration, so that dyspnœa may be produced.

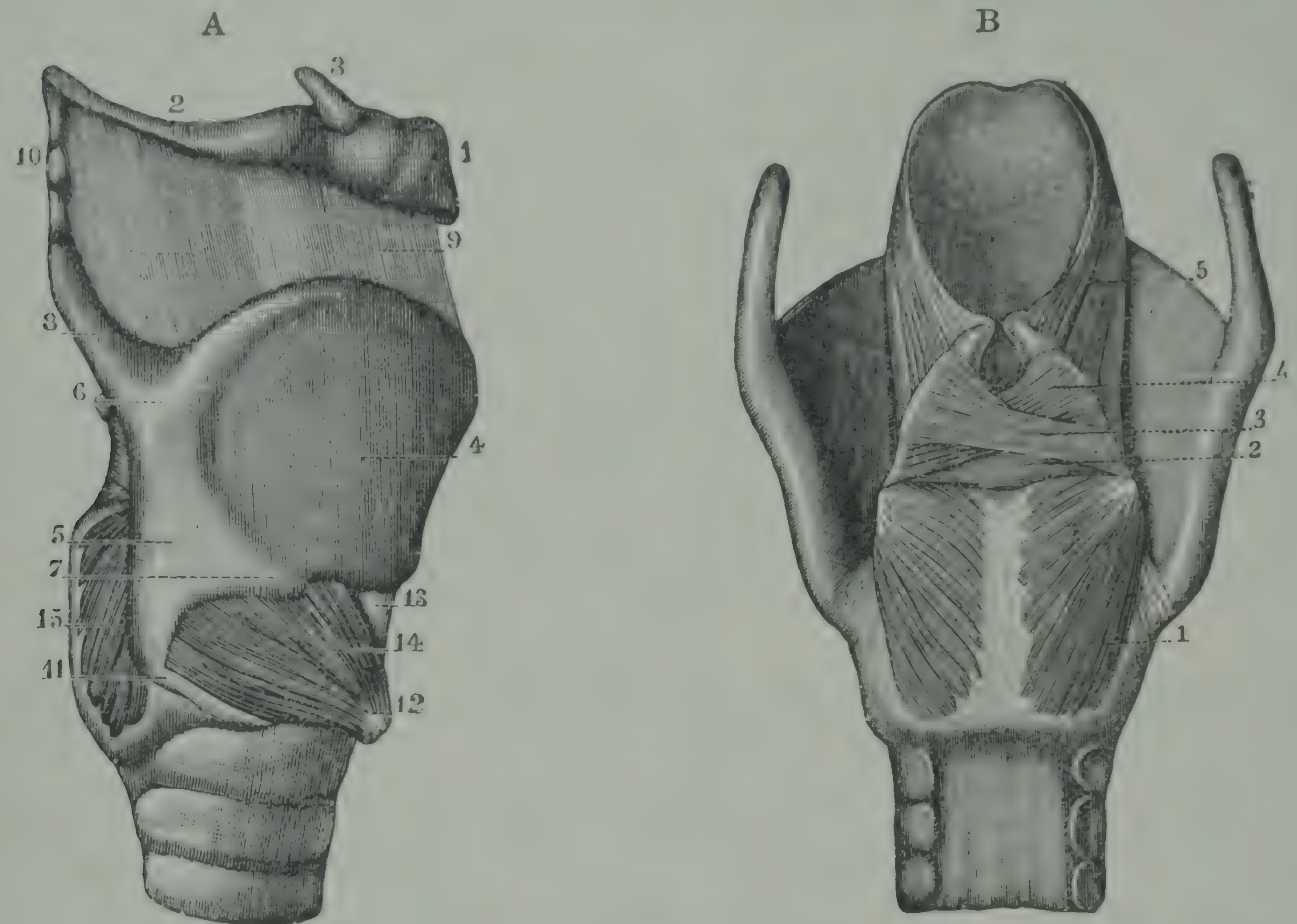


FIG. 314. Muscles of the Larynx. (SAPPEY.)

A, as shown in a view of the larynx from the right side.

1. Hyoid bone.
2, 3. Its cornua.
4. Right ala of thyroid cartilage.
5. Posterior part of the same separated by oblique line from anterior part.
6, 7. Superior and inferior tubercles at ends of oblique line.
8. Upper cornu of thyroid.
9. Thyro-hyoid ligament.
10. Cartilago triticea.
11. Lower cornu of thyroid, articulating with the cricoid.
12. Anterior part of cricoid.
13. Crico-thyroid membrane.
14. Crico-thyroid muscle.
15. Posterior crico-arytenoid muscle, partly hidden by thyroid cartilage.

B, as seen in a view of the larynx from behind.

1. Posterior crico-arytenoid.
2. Arytenoid muscle.
3, 4. Oblique fibres passing around the edge of the arytenoid cartilage to join the thyro-arytenoid, and to form the aryteno-epiglottic, 5.

(3) The *lateral crico-arytenoid* muscle arises from the upper border of the cricoid cartilage and passes backwards to be inserted into the muscular process of the arytenoid cartilage. These muscles when they contract pull the muscular process of the arytenoid cartilage forwards and downwards, thus approximating the vocal cords at their posterior ends and antagonising the action of the posterior crico-arytenoid muscles.

(4) The *arytenoid* muscles consist of transverse fibres, some of which decussate, uniting the posterior surface of the two arytenoid cartilages. When they contract they draw the arytenoid cartilages together.

(5) The *thyro-arytenoid* muscles consist of two portions. The outer fibres rise in front from the thyroid cartilage and pass backwards to be inserted into the lateral border and the muscular process of the arytenoid cartilage. Some of the fibres pass obliquely upwards

towards the aryteno-epiglottidean folds. These are often spoken of as a separate muscle, the *thyro-epiglottidean*. By their action they tend to draw the arytenoid cartilages forwards and to relax the vocal cords. The upper fibres may also assist in depressing the epiglottis. The inner fibres are called the *musculus vocalis*. They arise from the lower half of the angle of the thyroid cartilage, and, passing backwards in the vocal cords, are attached to the vocal processes and to the adjacent parts of the outer surfaces of the arytenoid cartilages. Many fibres do not run the whole distance, but end in an attachment to some part of the vocal cord. Although their action must be to draw the arytenoid cartilages forwards, yet, since they are contained in the vibrating portion of the vocal cords, they cannot by their contraction relax these cords. It is probable that they play a great part in determining the tension of the vocal cords after these have been put on the stretch by the action of the crico-thyroid muscles. They may possibly act as a sort of fine adjustment of the tension, the coarse adjustment being represented by the crico-thyroids.

The Production of Voice. In order to study the changes in the larynx which are associated with voice production, we must make use of the *laryngoscope*. The principle of this instrument is very simple. A large concave mirror with a central aperture is fixed before one eye of the observer, sitting in front of the patient or person to be observed. The latter is directed to throw his head slightly backwards and to open his mouth. In order to keep the tongue out of the way the patient is made to hold the end of it by means of a towel. The mirror is then so arranged as to reflect light from a lamp into the cavity

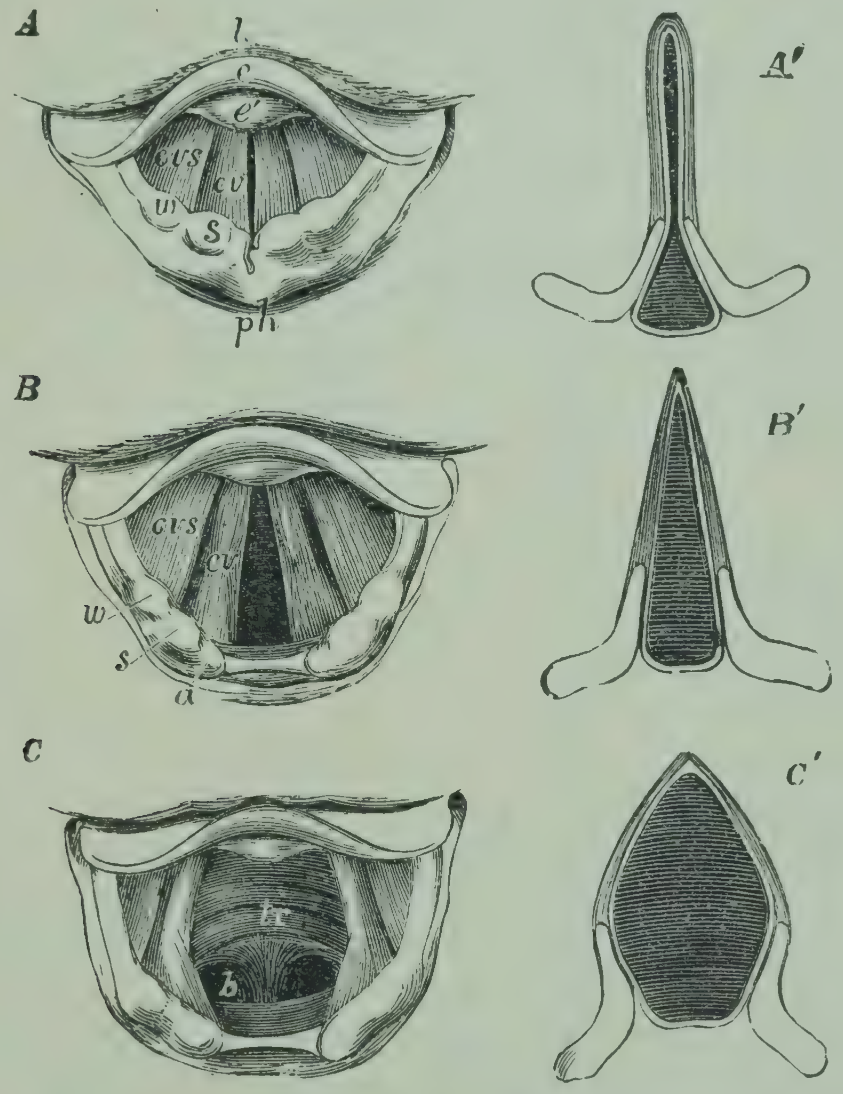


Fig. 315. Three Laryngoscopic Views of the Superior Aperture of the Larynx and surrounding parts on different states of the Glottis during life. (From Czermak.)

A, the glottis during the emission of a high note in singing.
B, in easy or quiet inhalation of air.
C, in the state of widest possible dilatation, as in inhaling a very deep breath.

A′, B′, C′ show in horizontal sections of the glottis the position of the vocal ligaments and arytenoid cartilages in the three states represented in the other Figures. Letters indicate as follows: *l*, base of tongue; *e*, upper free part of epiglottis; *e′*, tubercle of epiglottis; *ph*, part of anterior wall of pharynx behind the larynx; in the margin of the aryteno-epiglottidean fold *w*, the swelling caused by the cuneiform cartilage; *s*, that of the corniculum; *a*, tip of arytenoid cartilages; *cv*, true vocal cords; *cvs*, false vocal cords; between them the ventricle of the larynx.

In *C*, *tr* is placed on the anterior wall of the trachea, and *b* indicates the commencement of the two bronchi.

of the mouth. A small mirror fixed in a handle is then warmed, so as to prevent the condensation of the patient's breath, and passed to the back of the mouth until it rests upon and slightly raises the base of the uvula. By this mirror the light reflected into the mouth from the large mirror is again reflected down on to the larynx, and a reflection of the larynx and trachea is seen in the mirror.

By laryngoscopic examination we can see the base of the tongue, behind which is the outline of the epiglottis. Behind this again in the middle line are seen the two vocal cords, white and shining (Fig. 315, A). The cords appear to approximate posteriorly; between them is a narrow chink, the diameter of which varies with each respiration, being wider during inspiration. On each side of the true vocal cords are seen the pink false vocal cords. In some cases the rings of the trachea, and even the bifurcation of the trachea itself (Fig. 315, C) may be seen in the interval between the vocal cords.

In order that the vocal cords may be set into vibration, they must be put into a state of tension, and the aperture of the glottis narrowed so as to afford resistance to the current of air. In the dead larynx it is possible to produce sounds by forcing air from bellows through the trachea, after the vocal cords have been put on the stretch by pulling the arytenoid cartilages backwards. By experimenting on patients on whom tracheotomy has been performed, it has been found that the pressure of air in the trachea, necessary to cause production of voice, is, for a tone of ordinary loudness and pitch, between 140 and 240 mm. of water, and with loud shouting the pressure rises to as much as 945 mm. of water. This pressure is furnished by the contraction of the expiratory muscles, *i.e.* of the abdomen and of the thorax. Since the pitch of the note produced rises with increasing force of the blast, while the tension of the cords remains constant, it is evident that, in the act of "swelling" on a note, the increased pressure necessary for the crescendo must be associated with diminishing tension of the cords. It is the failure to secure this muscular relaxation that so often causes a vocalist to sing sharp when swelling on any given note.

The voice, like the sound produced on any musical instrument, may vary either in pitch, loudness, or in quality. The range of any individual voice is generally about two octaves. The pitch of the voice usually employed is determined chiefly by the length and mass of the vocal cords. Thus in children the voice is high-pitched. Before and at puberty there is a considerable development in the size of the larynx in both sexes. This is especially marked in the male, and accounts for the sudden drop in pitch ("breaking") of the voice. In the female the greater size of the larynx is chiefly perceptible in the increase in fulness and richness of the voice which occurs at this age. The ordinary human range, from deepest bass to highest soprano, is about four octaves. In singing, the voice may be produced in various ways, *i.e.*, in different registers. Thus we distinguish the chest register, the middle register and the head register. The deeper notes of any individual voice are always produced in the chest register. Observation of the vocal cord shows that when producing such notes the glottis forms an elongated slit, all the muscles which close the glottis and increase the tension of the cords being in action. The vocal cords are relatively thick and broad and can be seen to vibrate over their whole extent. When singing with the head voice, the vibrations of the cord are apparently confined to their inner margins; the aperture of the glottis is wider in front than behind, so that more air escapes during phonation by this method than in the production of the chest voice.✓

In order to change the pitch of the note the following means are probably employed in the larynx:

(1) Alteration in the tension of the vocal cords.

(2) Alteration in the length of the part of the vocal cords that is free to vibrate, which can be accomplished by the approximation of the arytenoid cartilage to one another, or by their approximation to the thyroid cartilage.

(3) The alteration in the shape of the vocal cords, which is determined by the activity of the different portions of the internal thyro-arytenoid muscles.

(4) The varying pressure of the blast of air passing through the glottis.

The loudness of the tone produced is practically proportional to the force of the blast of air employed. The quality or timbre of the voice depends not so much on the vocal cords as on the accessory resonating apparatus, represented by the trachea and chest and by the cavities of the mouth and nose. The greater part of the education involved in voice training is directed to the modification of the shape of the mouth cavity, so as to secure the greatest possible fulness, *i.e.* richness in overtones, of the tone produced in the larynx.

The Acoustics of Speech. The sounds employed in speech, viz. vowels and consonants, are produced by modifying the laryngeal tones by changes in the shape of the throat, mouth and nasal cavities. Vowel sounds are continuous, whereas the consonants are produced by interruptions, more or less complete, of the outflowing air in different situations. In whispering speech there is no phonation at all, but the sound is produced by the issue of a blast of air through a narrow opening between the lips, between the tongue and soft palate, or between the tongue and the teeth.

VOWELS. During phonation the expired air is forced under pressure between the tense vocal cords thus setting them into vibration. A series of fairly regular puffs of air thus pass up the throat to the mouth and nose. This series of puffs constitutes a musical tone which is very rich in harmonics. Certain of these harmonics may be emphasised by the resonant vibrations of the throat, mouth or nose. The various vowels are produced by the variations in the pitches of these accentuated harmonics. Thus on singing ū (as in tool) with a fundamental pitch of 100 v.p. sec. the third harmonic (400 v.p. sec.) and the seventh harmonic (800 v.p. sec.) are accentuated. If, on the other hand, the third and the twenty-first harmonic be accentuated the vowel becomes i, as in will.

The characteristic frequencies of components of the vowel sounds when the fundamental is 100 v.p. sec. are given in the table below.

Series.	Speech Sound.		Low Frequency.	High Frequency.
	ū	mule	**400**	800
	ō	mole	**500**	850
I.	au	maul	**600**	950
	ŏ	moll	**700**	1150
	ar	marle	**825**	1200
	ē	feat	375	**2400**
II.	i	fit	450	**2200**
	ā	fate	550	**2100**
	ă	fat	750	**1800**
III.	ō	foal	**500**	**800**
	ur or er	furl	**500**	**1500**
	ĕ	fell	**550**	**1900**
	ā	fail	**550**	**2100**

In the first series the low frequency resonance is loud and the high frequency one is quiet, and they rise in pitch together. In the second series the converse is found to be the case, for the low frequency resonance is quiet and the high frequency resonance is loud; moreover, as one rises, the other falls. In series III., which may be called the transitional series, since it connects series I and II., the low frequency resonance remains approximately constant in pitch, where the high frequency resonance changes greatly.

If these three series be plotted on a graph with the low frequencies as abscissæ and the high frequencies as ordinates a triangular figure is obtained.

There has been much controversy as to whether the pitch of these harmonic components changes at all with the pitch of the voice, or varies in different individuals. Some said that they did not change, others that their pitch kept in constant ratio with the pitch of the note sung; if the note doubled in pitch, so also did those of the vowel components. Several methods have been employed for investigating this point:

(1) By recording the vibrations emitted by the voice by means of the manometric flame.

(2) By analysis of the vibrations recorded by means of a gramophone.

(3) By measuring the intensity of vibration of series of resonators. All the above methods show that there must be some change, even if it is slight.

(4) By running a gramophone record of a bass voice at an increased speed so that the notes were those of a treble. If now the pitches of note and vowels were in constant ratio, the quality of the vowels should not change when the speed is thus increased. Experiment shows that the words are greatly altered, losing their O's and A's and taking E and I instead. This shows clearly that the change in pitch of the vowels is not nearly as great as that of the note sung with them. We must conclude therefore that neither those who say there is no change, nor those who say there is constant ratio, are right, but that the truth lies between the two extremes. Diphthongs are produced by changing the form of the mouth cavity from that of one vowel sound to another, thus AI (the English I) = ah-ee run together and abbreviated.

CONSONANTS. These are sounds produced by a sudden check being placed in course of the expiratory blast of air by closure of some part of the pharynx or mouth. They are classified into labials, dentals, or gutturals, according as the check takes place at the lips, between teeth and tongue, or between back of tongue and soft palate. Each of these again can be divided into soft and hard consonants as they are accompanied or not by phonation. Thus when we pronounce D the production of the laryngeal sounds goes on during the check of the sound produced at the teeth, whereas with T there is an absolute interruption of phonation during the pronunciation of the consonant. It is thus practically impossible to make any marked difference between hard and soft consonants when whispering.

In the production of nasal sounds, such as NG, the mechanism is the same as for the production of B, D, G, except that the posterior opening of the nares is not kept shut by the soft palate, so that part of the sound comes continually through the nasal passages, when it acquires a peculiar resonance. These sounds are on this account often spoken of as "resonants." The aspirates are produced by the passage of a simple blast of air through a narrow opening which may be at the throat as in H, between tongue and teeth as in TH, or between lips and teeth as in PH or F.

The vibratives, such as R, are formed by placing the tip of the tongue, or the uvula, or the lips, in the path of the blast of air so that they are set into vibration by the blast. The vibrative R employed in English is entirely due to the tongue.

The sibilants, which may be voiceless, as in S, or accompanied with phonation, as in Z, consist of continuous noises produced by a narrowing of the path of the air between the tongue and the hard palate. They are therefore similar in production to the aspirates. In the production of the sound L, the tongue is applied by its edge to the alveolar process of the upper jaw, so that the air or voice escapes by two small apertures in the region of the first molar and between the inner side of the cheek and the teeth. The acoustic characters of these various consonants are still but imperfectly studied.

Much work has recently been done, notably at the Bell and H.M.V. laboratories, to determine the conditions necessary for speech to be intelligible. In the first place the sounds must have adequate intensity. In the second place an adequate band of frequencies must be present. Thus on cutting out low frequency tones from 16 to 400, no loss of intelligibility results. On cutting off from 16 to 500 a 2 per cent. loss results. On cutting out from 16 to 600 a 10 per cent. loss results. With regard to high frequencies, on removing those above 4,000 to 5,000, little or no loss is found. On cutting off those above 3,000, there is a 10 per cent. loss in intelligibility. Thus for good speech, frequencies of 500 to 4,000 should be transmitted to the ears. For adequate speech, probably 600 to 3,000 would be sufficient. In the third place, " hang-over " or reverberation must not exceed about half a second. When a sound is generated in an empty, undraped hall, it continues for some time to reflect to and fro with gradually decreasing intensity. This phenomenon is called reverberation. Knudson carried out investigations in a room the reverberation of which could be altered by lining it with various thicknesses of felt. He found a decrease of intelligibility of speech of 7 per cent. for each second of reverberation time. A room with a reverberation time of half a second should cause a 3½ per cent. decrease of intelligibility, which is a hardly noticeable amount.

The Neurology of Speech. The intricate act of speech involves the employment of three different mechanisms : (*a*) The lungs, (*b*) the larynx, and (*c*) the cavities of the pharynx, mouth and nose. The lungs develop the necessary air pressure. The vocal cords of the larynx set the air into vibration with the required pitch. The cavities above, *e.g.* the mouth, modify the intensity of the tone and the relative intensities of its overtones so as to produce words.

THE LUNGS. For purposes of respiration the lungs are alternately filled and emptied by the synchronous action of the muscles of the chest wall and of the diaphragm. In addition, the muscles of the abdomen and arms and neck may be called into play when occasion demands more thorough ventilation. These muscles are innervated by nerve fibres which originate in the cells of the anterior horn of grey matter in the spinal cord. These same muscles are synchronously brought into play for purposes of phonation by voluntary impulses which, for a time, replace the normal respiratory rhythm. In general, inspiration is quickened, the vocal cords are separated, and air pours into the lungs. The vocal cords are now brought together and put under tension. Expiration now begins, and air commences to pass between the cords ; these vibrate, producing sounds which are modified into speech. Expiration is prolonged until we have completed the sentence. The expiratory pressure even in quiet speech is above that during ordinary respiration, while in shouting not only the ordinary expiratory muscles are forcefully contracted, but the accessory muscles of expiration (*e.g.* the abdominal) are freely made use of as well. Considerable muscular effort is required for loud speech or song. So great may it be that the individual becomes breathless as he does when performing other kinds of severe muscular effort.

THE MUSCLES OF THE MOUTH, Etc. There are six points at least at which modification in diameter of the air cavities above the vocal cords may be produced : (*a*) The false cords, innervated by the recurrent laryngeal branch of the vagus nerve ; (*b*) the pharynx at the level of the root of the tongue, by the pharyngeal branch of the vagus ; (*c*) the buccal cavity near its posterior part, by the elevation or depression of the soft palate, supplied by the spinal accessory nerve ; (*d*) the buccal cavity near its middle, by the elevation or depression of the dorsum of the tongue, innervated by the hypoglossal nerve ; (*e*) the buccal cavity near its front, by the approach of the tip of the tongue to the upper teeth ; (*f*) the buccal cavity at its anterior outlet, by the approximation of the lips (facial nerve).

In addition, the mouth cavity may be varied in size by the change in angle of upper and lower jaw (trigeminal nerve), and the nasal cavity may be allowed to communicate with the pharynx at its posterior outlet (spinal accessory nerve).

We may summarise these nervous controls by saying that the seventh, tenth and twelfth cranial nerves play the principal part in controlling this mechanism of speech.

THE CORTICAL CENTRES OF SPEECH. At one time it was thought that there was a circumscribed motor area for speech. It was called Broca's area after its discoverer, and occupies the posterior part of the inferior frontal gyrus. In right-handed people it is found on the left side of the brain only, and *vice versâ*. Destruction of this area undoubtedly causes dumbness, but histological evidence shows that it is an association area only. After destruction of this area the lungs can still be voluntarily ventilated. The cords can still be approximated and set into vibration, the false cords, pharynx, tongue, and lips can all be separately moved, but the complex correlation of all these separate motor functions is no longer possible. We may say, then, that, as the result of external stimuli, impulses reach Broca's association area, and by it they are distributed in suitable combination to the motor cortex of lips, tongue, cords, diaphragm, etc. The impulses travel down the axons of the Betz cells thus innervated, through the corona radiata, internal capsule, mid-brain, and pons, to reach either directly or through intermediate connections the motor nuclei of the seventh, tenth, and twelfth cranial nerves and the anterior horn cells of the spinal segments involved in the act.

Broca's area is by no means the only important cortical association area involved in speech ; others are found near both the auditory and visual cortex. ✓

APHASIA. It has been usual to divide the disorders of speech, known as aphasia, into the following groups : (1) motor aphasia of Broca ; (2) sensory aphasia of Wernicke ; (3) alexia ; (4) anarthria. A more fundamental idea of the processes involved will, however, be obtained if it is remembered that there are two ways in which the necessary ingoing impressions are received, namely, by reading written or printed words and by hearing spoken words. There are, correspondingly, two different mechanisms by which language may be conveyed, namely, by writing and by speaking. Consideration of these facts would suggest that there may be eight different possible defects: (1) vision may be perfect and yet there may be complete inability in understanding written or printed words (alexia); (2) hearing may be perfect and yet the man may be incapable of understanding what is said to him (this is called sensory aphasia of Wernicke); (3) he may be able to use his hands and to draw with a pencil and yet be unable to write (agraphia); (4) he may be able to use his lips and to make noises with his throat and yet be unable to speak (anarthria or motor aphasia of Broca). Each of the above defects concerns one particular mechanism alone. Other possible defects may concern two mechanisms, one ingoing and one outgoing. Thus (5) there may be ability to read with inability to copy what is read, or (6) ability to read but inability to read aloud, or (7) ability to understand what is spoken but inability to write it down to dictation, or (8) ability to understand what is spoken and yet inability to repeat it.

The neurological basis of language must be regarded as co-extensive with the sensory centres, and with the whole region of lower association. We might speak of auditory and visual word-centres as located in the visuopsychic and auditory psychic centres. There is probably, however, no word, still less a collection of words, expressing an idea which does not involve the activity of practically all parts of the cerebral cortex. As Bolton points out, "a word, such as 'mouse,' at once sets in effect processes of association which pass to every projection sphere with the solitary exception of the gustatory, and even this may be aroused in a person who has eaten a fried mouse in the hope of thereby recovering from an attack of whooping-cough." The amount of impairment of intelligence will vary in different cases according to the extent of the lesion. Thus, softening affecting the occipital lobe may, with hemianopia, cause "word-blindness" or "alexia," a loss of power of appreciating the meaning of written words. In most individuals, and certainly in the uneducated, this power may be cut out altogether without interfering considerably with the mental powers. On the other hand, from babyhood upwards we have learnt the meaning of words and their grouping by auditory impressions. If the whole of the auditory associations be destroyed by an extensive lesion in the first and second temporal convolutions, the resulting loss of word appreciation, sensory aphasia, will be attended with great diminution of mental powers. It must be remembered that the area of Wernicke is not a sensory centre, but a centre of association between the various sense impressions, especially those of hearing and sight. It may therefore be spoken of as an intellectual centre. Pure motor aphasia, if it exists, is always anarthria and is due to a lesion in the lenticulate zone, *i.e.* in the lenticular nucleus and its neighbourhood, in the anterior part and the genu of the internal capsule and possibly in the external capsule.

THE FUNCTION OF EXPRESSION. Although, since it is extremely exact, we rightly give precedence to speech as a means of conveying our ideas to one another, other modes of expression are of considerable interest since they probably once played a more prominent part than they do at the present day. No attempt will be made here to analyse them in detail. They may be divided anatomically into facial, manual, corporal, and pedal. The first comprises all those movements of expression of the superficial muscles of the face, the lips, the jaw, tongue, etc., which are called into play. Thus, corrugation of the superficial muscles of the forehead indicates surprise or worry; if the eyebrows are raised at the same time it indicates surprise; if they are lowered it indicates anger. The orbicularis muscles of the eyes when relaxed show astonishment or interest, and when contracted indicate boredom. Tremor of the lids expresses extreme mental agitation. The lips may be pursed up to show doubt, spread out in a smile, or separated in a grin.

Manual expressions are not so easily analysed as the above, but may be divided arbitrarily under a time, space, and intensity distribution. For example, the placing of the first finger on the lips to give the order for silence;

patting the hands quietly together to signify approval; the slow raising and lowering of the right hand, palm downwards, to signify go slowly or stop. They are largely matters of convention, and some races use them much more than others. The conductor of an orchestra will show many examples of the expressions which come under this category.

As examples of corporal expressions, we have the leaning forward in a chair when interested, or leaning back when bored, or crouching as if to spring when in disagreement with what is being said, or when angered with the speaker, and as instances of pedal expressions, the shuffling of the feet when the mind is anxious, or the stamping of the foot when angry.

All the above are under the control of the will. In addition, there are the emotional symptoms controlled by the sympathetic system which are not under voluntary control—namely, the erection of the hair (the hair " stands on end "), the dilatation of the pupil, the activity of the sweat glands, stimulation of the lachrymal glands expressing sorrow, and the cessation of the secretion of the buccal glands when the mouth becomes dry accompanying an anxious state of mind.

For Reference

HARVEY FLETCHER (1929). Speech and Hearing. (London: Macmillan & Co.)

LASHLEY (1923). The Integrative Functions of the Cerebral Cortex. *Physiol. Rev.* **13,** 1.

KERRIDGE (1938). Recent Advances in Knowledge Concerning Hearing and Speech. *Physiol. Rev.,* **18,** 59.

CURRY (1939). The Mechanism of the Human Voice. (Churchill.)

CHAPTER XXVII

CUTANEOUS SENSATIONS

The skin, being the outermost layer of the body, represents the tissue or organ by which the organism is brought into relationship with its environment. In the widest sense of the term the skin is protective. This function it discharges by virtue not only of its physical properties, but also of its rich endowment with sense organs, by means of which the intracorporeal events can be correlated with those occurring outside and immediately affecting the organism.

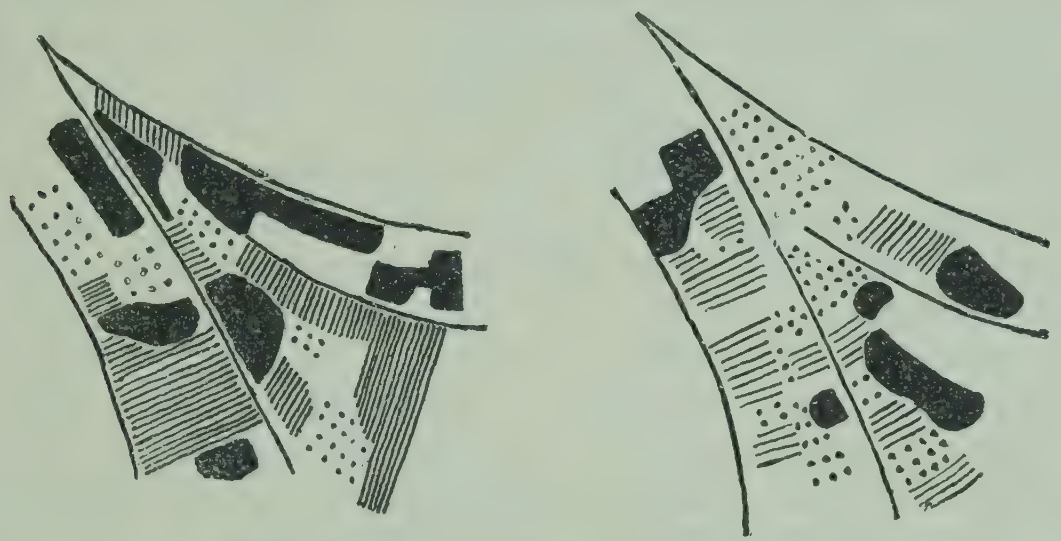
We are accustomed to distinguish several qualities of sensation among those having their origin in the skin, the chief of which are the sense of touch, including that of discrimination, the sense of pain and the sense of temperature. The very different qualities of sensation included under these three classes suggest that there may be a special mechanism, or class of mechanism, for each sense, and a careful investigation of the sensory qualities of the skin surface bears out this idea. Isolated stimulation of minute areas on the skin does not excite all the sensations together, but only a sense of touch or of pain, or a sense of cold or warmth. We are therefore justified in dealing with each of these sensations separately.

THE TEMPERATURE SENSE

By means of the skin we can appreciate when a body coming in contact with it is either cold or warm. If the body is at the same temperature as the skin, as a rule no sensation of temperature is excited. It was formerly thought that the sensations both of heat and cold were determined by the excitation of one and the same end organ. The warming of this end organ would produce a sensation of warmth, while a diminution of its temperature would produce the sensation of cold. Careful investigations of the distribution of the temperature sense by Blix and Donaldson has shown that this opinion cannot be maintained. If a small object warmed to a few degrees above the temperature of the skin be moved over any part of the surface of the body, *e.g.* the back of the hand, it is found that the warmth of the instrument is not equally appreciable at all parts of the surface of the skin. At some points the sensation of warmth will be very pronounced, but between these points the sensation of warmth may be entirely wanting and the instrument may be judged to be of the same temperature as the hand itself. In this way a series of " warm points " may be mapped out. On now cooling the instrument a few degrees below the temperature of the surface of the body and then moving it over the surface in the same way, it will be found again that the coolness of the instrument is appreciated only at certain points ; these can be regarded as " cold points " and as containing the nerve endings by the excitation of which the sensation of cold is produced. If the warm points be pricked out in red ink and the cold points in blue ink, it will be seen that they do not in any way correspond.

A convenient instrument for this purpose is the one invented by Miescher, consisting of two tubes cemented together and communicating at a small flattened extremity, which is applied to the surface of the skin ; water can be led through the tubes at any desired temperature, which is read off by a thermometer placed within the tube. Having mapped out the warm spots, it may be shown that they are excitable by means of mechanical or electrical stimuli and that the sensation produced is the same as if they had been excited by their adequate stimulus, viz. a rise of temperature.

Exact Location of the spots is rendered difficult by the irradiation of the sensation produced, so that it is difficult to refer the sensation of warmth or cold definitely to the point stimulated. An investigation of the topography of these warm and cold spots shows that the apparatus for the appreciation of cold is much more extensively distributed over the body than that for the appreciation of warmth, as is evidenced from the diagram (Fig. 316) giving the topographic distribution of the cold and warm sense organs on the palm of the hand. Waterston (1923) claims that in the case of both heat and cold, the topographic distribution of the spots alters continually. In skin suffering from a mild erythema, all parts respond to warmth, that is, all signs of a punctate arrangement have gone. He concludes that all parts of the skin can respond to both heat and cold, the punctate arrangements found at a single examination being due to some parts being more active than others. The skin capillaries are stated to show a similar change from time to time in the quantity of blood which they transmit. The temperature sense is best marked in the following regions of the body : the nipples, chest, nose, the anterior surface of the upper arm and the anterior surface of the fore-arm, and the surface of the abdomen. It is much less marked on the exposed parts of the body, such as the face and hands, and is but slight in the mucous membranes. Thus it is possible to drink hot fluid, such as tea, at a temperature which would be painful to the hand, and still more to other parts of the body. The scalp is also very insensitive to changes of temperature. The acuteness of the temperature sense varies considerably with the condition of the skin and with the previous stimulation of the sense organs. The sense is most acute at about ordinary skin temperature, *i.e.* between 27° and 32° C. At this temperature the skin can appreciate a difference of $\frac{1}{8}$° C. When the skin is very cold or very hot, the temperature sense is not nearly so delicate. This sense presents the phenomenon of adaptation in a marked degree. It is a familiar experience that, on coming from the external air on a cold day into a warm room, a sensation of warmth is experienced all over the body. In a few minutes this sensation wears off. On now leaving the room to go outside again, the sensation of cold is at once appreciated, to disappear in its turn after a few minutes. The effect of adaptation is still better shown by the experiment of taking three basins of water, *a*, *b*, and *c* ; *a* contains cold water, *b* tepid water, *c* hot water. The left hand is immersed in the cold water and the right hand in the hot water for a few minutes. On now placing both hands into the basin of tepid water, it feels hot to the left hand and cold to the right hand. Such experiences as this led Weber to the conclusion that the essential stimulus for the temperature sense was not the actual temperature to which the sense organs were subjected, but the fact of a change of temperature. He imagined that, while the temperature sense organs were being warmed, a sensation of warmth was produced, and when their temperature was being lowered, a sensation of cold. Such a theory,



Cold spots. Heat spots.

Fig. 316. Heat and Cold Spots on part of Palm of right Hand.

The sensitive spots are shaded, the black being more sensitive than the lined, and these than the dotted parts. The unshaded areas correspond to those parts where no special sensation was evoked. (Goldscheider.)

however, would not account for the fact that, above a certain temperature, water may feel warm and the feeling may continue so long as the skin continues to be stimulated. On a cold day the air may feel cold to the face and the feeling may last the whole time that the face is exposed. Moreover we have in the temperature sense, conditions which remind one of the after images which occur in the eye. If a cold object be pressed on the forehead and then removed, the sensation of cold lasts some little time after the object has been removed. In this case a sensation of cold is produced although the end organs are being gradually warmed up after the removal of the cool object. In order to account for these facts Hering, at a time when the differentiation of hot and cold spots had not yet been effected, suggested that the temperature sense organs could be regarded as having a zero point at which no sensation was produced. If their temperature was raised above this point, a sensation of warmth was produced, and *vice versâ*. The zero point, however, was not a fixed one, but could move upwards to a certain extent on prolonged exposure to high temperature, or downwards on prolonged exposure to a low temperature. In the light of the researches of Blix and Goldscheider, we should have to apply Hering's theory of a zero point to each of the temperature end organs separately.

A cold pencil passed over a warm spot evokes no sensation whatsoever. If, however, a pencil considerably warmer than the skin be passed over a cold spot, this may be excited so that the paradoxical result is produced of a sensation of cold as the result of stimulation by a warm body. It is a familiar fact that the immediate effect of entering a hot bath is very much the same as that of entering a cold bath, viz. a rise of blood pressure and contraction of the unstriated muscles of the skin and hair follicles with the production of "goose skin." It has been suggested that the distinctive quality of a sensation of *hot* as compared with that of *warm* is due to the simultaneous stimulation of cold spots. With regard to the identity of the sense organs for warmth and cold, J. P. Nafe's theory must be referred to, namely that the sensation of warmth is brought about by dilatation of the skin vessels, and that to some extent the sensation of cold is brought about by their contraction. He supports his theory by evidence among which we may notice (1) the sensation of flashes of warmth and cold when the arteries of the skin are punctured * ; (2) the striking proximity of blood vessels to cold points ; (3) the absence of blood vessels and the absence of cold and warmth sensations from the cornea ; (4) the similarity of the reaction times for the appreciation of warmth and cold, and the contraction and dilatation of the skin blood vessels.†

THE SENSE OF TOUCH

By means of the sense of touch we arrive at a conclusion as to the qualities, such as shape, texture, hardness, etc., of the bodies with which the skin is in contact. In this judgment, however, very many other sensations are involved besides those which can be regarded as strictly tactile. Thus, the hardness of an object signifies its resistance to deformation, besides its power of deforming the skin surface with which it is in contact ; the former quality, *i.e.* of resistance, is one which involves the muscular sense, since we judge of it by the extent to which we can move our muscles without causing any alteration of the surface of the object.

The tactile sensibility of the skin as a whole, like its temperature sensibility, was thought to have a punctate distribution. Recent work on the subject, by Waterston (1923), throws very great doubt on the validity of this idea.

RESPONSE TO DIFFERENT STIMULI. The adequate stimulus for the tactile nerve endings is not so much pressure, as deformation of surface. It appears to matter little whether the surface be deformed by pulling it or by pushing an instrument into it. The ineffectiveness of mere pressure is shown by dipping the finger into a vessel of mercury. The sensation of pressure is noted only at the point where the finger passes through the surface of the mercury, and this is the only part where there is an actual deformation of the skin, due to the sudden passage from the pressure of the mercury to the negligible pressure of the outside air. The tactile apparatus is smarter in its response than any other of the sense organs. On this account stimuli are still perceived as discrete, when they are repeated at a rhythm which would result in complete fusion in the case of any of the other sense organs. Thus if a bristle be attached to a tuning fork and allowed to press on the skin, the vibrations of the fork are perceived by the ear as a continuous sound and by the skin as a series of discontinuous taps. Faradic currents when applied to the skin can be perceived as separate when repeated at the rate of 130 per second. The sensations evoked by placing the finger against the edge of a cog-wheel do not become continuous until the wheel is revolving at such a rate that the stimulation on the skin by the serrations occurs at a greater rate than 500 or 600 per second. The tactile apparatus resembles all the other skin sense organs in showing adaptation. A stimulus after continuing for some time may become ineffective. We are usually entirely unaware of the stimulation of our skin by the pressure of the clothes, and even an unwonted stimulation, such as that of the mucous membrane of the mouth by a plate carrying artificial teeth, though almost unbearable during the first day, rapidly becomes less, and in a few days it is not perceived at all.

---

* BAZETT and MCGLONE. *Brain*, **51**, p. 18, 1928.

† See p. 1057, "Handbook of General Expt. Psychology." (Clark Univ. Press, 1934.)

In order to test the sensitiveness of touch we may use the method introduced by Hensen, viz. the bending of a glass-wool fibre. We can determine the pressure at which any given fibre will bend, and if we find by trial the fibre which just evokes sensation when pressed on the skin, we know exactly the force which we are applying to the skin. Von Frey employed hairs of different thickness for the same purpose (Fig. 317). The following represents the minimal excitability of the surface of different parts of the body when tested in this way:

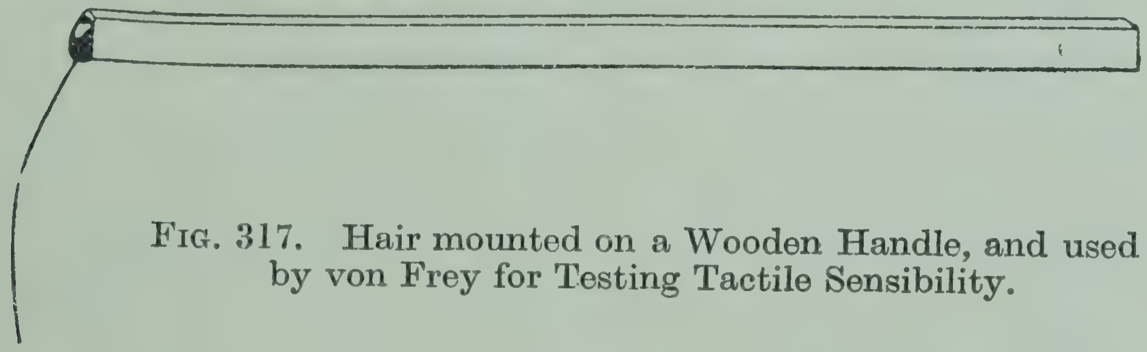


FIG. 317. Hair mounted on a Wooden Handle, and used by von Frey for Testing Tactile Sensibility.

	Grm. per sq. mm.
Tongue and nose	2
Lips	2·5
Finger-tip and forehead	3
Back of finger	5
Palm, arm, thigh	7
Fore-arm	8
Back of hand	12
Calf, shoulder	16
Abdomen	26
Outside of thigh	26
Shin and sole	28
Back of fore-arm	33
Loins	48

The sensitiveness of the sense organs in the skin is probably much greater than that of the nerve trunks themselves. Thus Tigerstedt found that the minimal mechanical stimulus necessary to excite the exposed nerve amounted to 0·2 grm. moving at 140 mm. per second. For the touch spots, von Frey found that 0·2 grm. moving at 0·17 mm. a second is an adequate stimulus.

In testing the sensibility of any surface, it is important to remember that the hairs themselves form very effective tactile organs. The touch spots are distributed in greatest profusion around hair follicles, and there is a rich plexus of nerve fibres round the root of each hair. A slight touch applied to the hair acts on these as on the long end of a lever, the hair being pivoted at the surface of the skin, so that pressure on the hair, increased five or more times in force, is transmitted to the hair follicle and the surrounding nerve endings. The actual sensibility of any part is therefore much diminished by removal of the hairs. On 9 sq. mm. of the skin, from which the hairs had been shaved, the minimal stimulus necessary to evoke a tactile sensation was found to be 36 mg., whereas on the same surface before it was shaved 2 mg. was effective.

WEBER'S LAW. The smallest increment or decrement of stimulus which determines a perceptible difference of sensation must, according to Weber's law, always bear the same ratio to the whole stimulus. In measuring such differences it is best to apply the stimulus successively to the same surface of the skin, rather than simultaneously to adjoining areas. The time interval between two successive stimuli should not be more than five seconds and the duration of the stimuli should be equal. Weber found that in the terminal phalanx of the finger the minimal perceptible difference was about one-thirtieth, but the ratio was not the same for all regions of the skin nor for all individuals. The following represents the liminal difference in various skin regions :—

Forehead, lips and cheeks	1/30th to 1/40th
Back of fore-arm, of leg and of thigh ; back of hand, and first and second phalanx of finger, etc.	1/10th to 1/20th
All parts of the foot, leg and thigh	more than 1/10th

THE SPATIAL QUALITY OF TOUCH. DISCRIMINATION. If any part of the skin be stimulated, the subject of the experiment can tell at once the exact situation of the excited spot. If two points be stimulated simultaneously, excitation is perceived as double, *i.e.* as proceeding from two points, provided the distance between the points exceeds a certain amount, varying in different parts of the body. The power of discrimination, *i.e.* of judging whether a stimulus is single or double, can be tested by arming the points of a pair of compasses with small pieces of cork and then seeing how far apart these must be when pressed on the skin in order that the stimulus may be perceived as double. The following Table represents this distance for various regions of the body :—

DISTANCE IN MM.

Skin region.	mm.
Tip of tongue	1·1
Volar surface of finger tip	2·3
Dorsum of third phalanx	6·8
Palm of hand	11·3
Back of hand	31·6
Back of neck	54·0
Middle of back, upper arm, and thigh	67·1

When touch spots are sought out for stimulation with the points of a compass, the distance at which the excitation is perceived as double is much diminished, as is shown by the following Table of distances for the touch spots in millimetres :

Skin region.	Distance of touch spots.
Volar side of finger tips	0·1
Palm of hand	0·1
Fore-arm (flexor side)	0·5
Upper arm	0·6
Back	0·4

The compass points are perceived to lie apart with a special distinctness when they are applied to touch spots lying on different lines which radiate from the hair follicles. The figures given in the first Table have no relation to touch spots, but show the average distance at which no excitation can be perceived as double.

The delicacy of discrimination of any part is largely associated with its mobility. Thus, in the arm the delicacy increases continuously from the shoulder to the finger tip. If the localising power for touch on the shoulder be taken as 100, that of the finger tips will be represented by 2,582. In the same way there is a continuous decrease of the distance of discrimination as we pass along the cheek from the ear to the lip, *i.e.* from the nonmobile to the mobile part. The power of discrimination is increased to a certain extent by practice and largely diminished by fatigue. Any factor which diminishes the tactile sensibility of the part, such as cold, will also diminish the power of discrimination.

Localisation of Touch. The fact that we can localise the point of stimulation shows that every tactile sensation derived from the surface of the body, besides the qualities of intensity and extensity, has also associated with it a characteristic quality dependent on its position. This localised quality of a tactile sensation was called by Lotze "local sign." Among psychologists there has been much discussion as to how far this local sign is an inborn attribute of the sensation of every point on the body surface, or how far it is acquired by experience and based on memory of movements and muscular impressions. In the retina we have a sense organ which, like the skin, possesses local sign but in far higher degree, the power of discrimination of the retina being three thousand times as great as that of the most sensitive part of the skin. Cases of congenital cataract occur in which the subjects have been blind from birth. By extraction of the cataract we can give such persons the power of sight. It is found that at first there is no power of localising visual impressions. The local sign is developed only in response to experience, by comparing simultaneous visual, tactile and motor sensations. By analogy we might ascribe the local sign of cutaneous sensations to a similar causation. Our study of the spinal animal has indeed given us a physical or histological conception of local sign. We know that stimulation of any part of the body evokes an appropriate reaction, the nature of which is determined by the central connections of the entering nerve fibres. A fibre entering at one segment must therefore come into relation with a different set of motor cells from those which are set into action by a fibre entering one segment lower down. Every nerve fibre from the skin will therefore have an appropriate complex of motor paths in functional connection with its central endings ; and when the activity of these reflex paths comes to be represented in consciousness, it is evident that the sensation derived from any one point must differ from that derived from any other point of the skin, by virtue of the differing motor events actually or potentially excited from the two points. In ascribing local sign to coincident muscular sensations and to the memory and experience of past movements, we are therefore giving but an imperfect explanation ; since the difference between the sensations from different parts, which is at the bottom of our powers of localisation, has its origin in the structure of the central nervous system itself and is present from the very beginning of the evolution of a reactive nervous system.

Projection of Touch. Since the alterations in the surface of the skin which give rise to tactile sensations are habitually caused by contact with external objects, we come to regard the sensations themselves, not as changes in the skin, but as qualities of the object which touch the skin, *i.e.* we project the sensation. The projection is, however, not so great as in the case of visual sensations. Cutaneous sensations we always consider as qualities of an object immediately affecting and altering the condition of ourselves, whereas the visual sensations are referred at once to objects lying right away from ourselves, so that we are not aware that any change has taken place in our bodies as a result of the entering of rays of light into the eye.

It is remarkable to what extent projection of touch sensation may occur.

Thus a surgeon actually lengthens his fingers by using a probe. When he is probing for dead bone he feels the grating of the bone, not at his finger tips, but he projects the sensation to the end of the probe. In the same way tactile sensations evoked by the contact of bodies with the insentient endings of hairs are referred to the ends of the hairs rather than to the hair follicles where the nerve impulses actually come into being.

The dependence of local sign on habitual experience is shown by the various tactile illusions, such as the well-known experiment of Aristotle. If with the eyes shut we cross the first and middle fingers and bring them in this position in contact with a pea, we should at once say that two peas lay under the fingers. This is especially marked if the pea be rolled between the fingers. The two sides of the fingers which come in contact with the pea usually touch two different objects, and these parts of the skin would have to be re-educated, *i.e.* their local sign would have to be changed in accordance with the changed conditions, before the pea would be perceived in its true state as single.

THE PAIN SENSE

When the pressure of a hard object on the skin is increased beyond that necessary to evoke a tactile sensation, at a certain pressure the quality of sensation changes and it becomes painful. For the evolution of the race as well as for the preservation of the individual this pain sense is all-important; it is the expression in consciousness of the reflexes of self-preservation which can be evoked in the spinal animal by stimuli which are nocuous, *i.e.* calculated to do actual damage to the tissues of the body. Thus when a sharp point is pressed on the skin, the sensation becomes painful just before the pressure is sufficient to cause penetration. The so-called trophic lesions, which occur in parts devoid of sensation, are determined for the most part by the lack of the pain sense and the consequent failure of the preservative reflexes of the part. It is remarkable that pain may result from changes in organs which are devoid of ordinary sensibility. Thus the intestine may be cut, sewn or handled without arousing any sensation whatsoever. A strong contraction of the muscular wall or increased distension of the gut will, however, evoke a griping pain. In the same way the ureters, which are normally devoid of sensation, can give rise to excruciating agony when they are contracted firmly on a retained calculus.

We are accustomed to distinguish many different qualities of pain, but on analysis it will be found that these qualities depend on the nature of the sense organ which is simultaneously stimulated. Thus a burning pain denotes simultaneous stimulation of the pain sense and of the nerve endings to the warm spots. A throbbing pain results when the vessels of the part are dilated and the part is tense with effused lymph, so that each pulse of the vessels causes an exacerbation of the painful stimulation and perhaps also stimulation of the tactile end organs.

The sense of pain has often been ascribed to over-maximal stimulation of any form of sensory nerve. Although it is true that over-stimulation of the auditory or optic nerve by a loud sound or a bright light may be extremely unpleasant, the sensations evoked do not partake of the characters of painful sensations such as would be produced by pricking or burning the skin. Moreover a careful investigation of the sensory points on the skin brings out the fact that there are, besides the tactile and temperature spots, other spots from which only painful sensations can be evoked. We have seen already that over-stimulation of a touch spot does not, as a matter of fact, cause pain. The pain spots which are distributed among the touch and temperature spots are insensitive to a low grade of stimulus. As the strength of the stimulus is increased, a point is suddenly reached at which the sensation evoked is painful. Moreover, in parts of the body, tactile and temperature sense are entirely wanting, though painful impressions can be easily evoked. The best example of this is seen in the cornea, minimal stimulation of which evokes pain, but nothing which can be regarded as a tactile sensation. The specific quality of pain sensation is shown moreover by the fact that in many cases of disease the sense of pain may be abolished without the sense of touch. Such a patient is said to suffer from analgesia, but not from anæsthesia. When an analgesic part is pricked the patient can say that he is pricked, but has no objection to any amount of repetition of the stimulus, since the sensation is entirely devoid of painful character. In the case of the skin, pain is found to differ somewhat from the other skin sensations in being much more uniformly distributed, more difficult to locate accurately, and more hardy. Thus while most sense organs are rendered less sensitive by cutting off blood supply, pain at first reacts more violently.

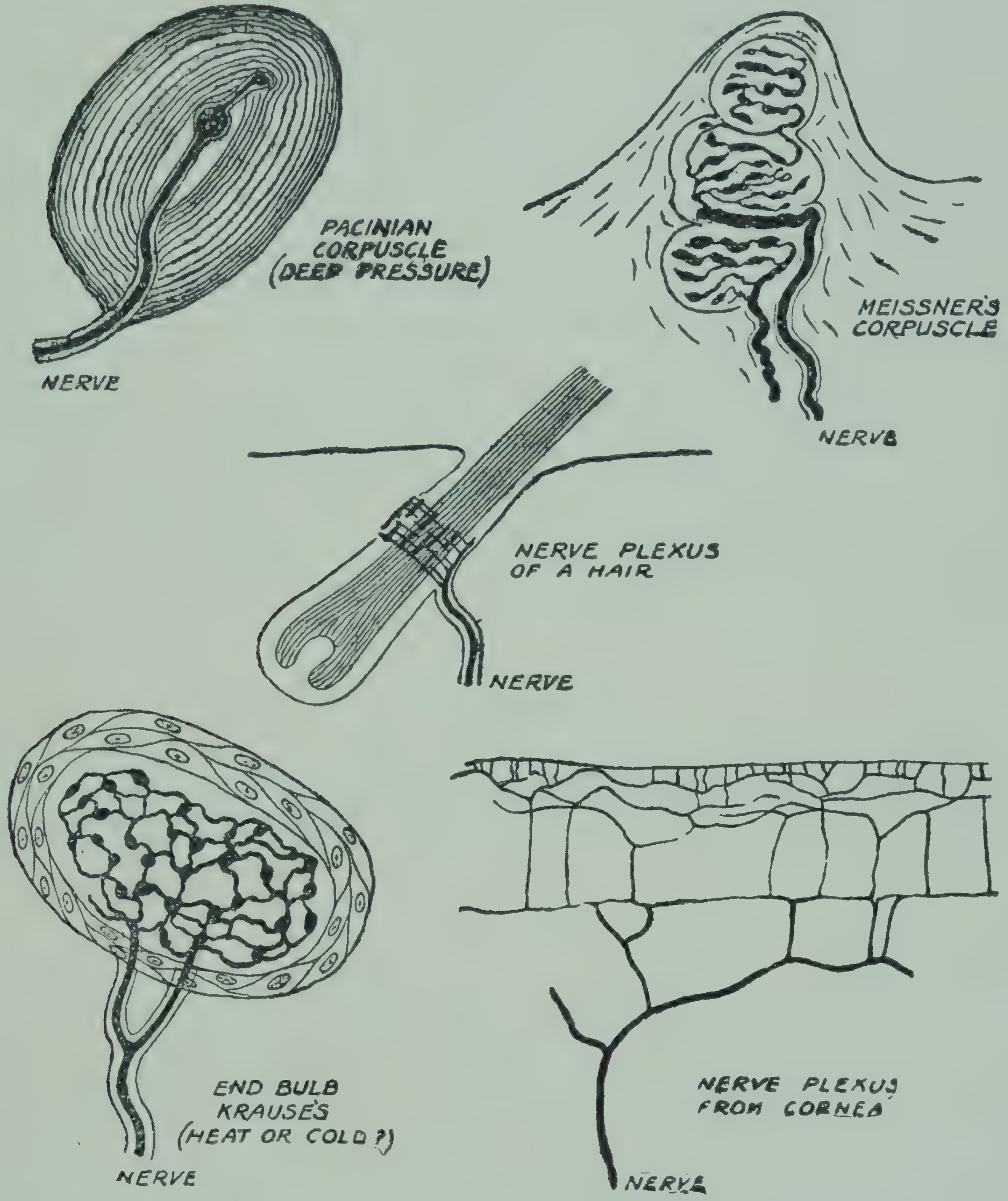


FIG. 318. Skin End Organs and the Sensations which they arouse. It has been supposed that corpuscles of the type of Meissner's corpuscle are for touch, and that nerve plexuses, such as that illustrated for the cornea, subserve pain sense.

Protopathic and Epicritic Sensations. As the result of experiments by Head and Rivers, in which a sensory nerve (the radial) in the arm was cut, and the returning sensibility followed during regeneration, it was thought that the sensations could be divided into three groups—epicritic, protopathic and deep. Deep sensations remained after the nerve had been cut, and consisted of pressure sense and pressure pain. Protopathic sensibility was the first to return and consisted of a vague pain sense and a sensation to hot objects above 38° C. and cold ones below 24° C. Epicritic sensation, the last to return, gave critical discrimination and localisation and differentiation of temperatures between 25° C. and 37° C.

But recent workers, Trotter and Davies (1909), and Boring (1915), who have repeated Head's and Rivers' experiment, and Carr (1916), who subjected Head's and River's own data to analysis, agree in the conclusion that the distinction into protopathic and epicritic is purely artificial and must be finally abandoned.

The Histological Character of the Elements Involved in Cutaneous Sensations

A very large number of different forms of sensory nerve endings have been described in relation to the skin. Their exact allocation among the different cutaneous senses presents considerable difficulties.

As regards touch, two kinds of elements are probably involved. In the first place, the most sensitive tactile apparatus are the follicles of the short hairs. Around these follicles we find a sheaf of nerve fibres, some of which end in the hair papillæ and others form a ring near the level of the openings of the sebaceous glands. The other tactile end organ is, according to Waterston (1932), the nerve plexus which terminates in the epithelial cells of the more superficial layers of the epidermis. The evidence he advances for this view is that on cutting away the superficial layers with a razor the sensation of touch is impaired or removed, leaving the sensation of pain unaffected. Other end organs, which are supposed to be stimulated by changes of pressure and therefore to be tactile, are the organs of Ruffini which occur in the papillæ of the palm and fingers, and, lying more deeply, the elastic tissue spindles as well as the Golgi corpuscles and the Pacinian corpuscles in the subcutaneous tissue.

As regards pain, we know that in the cornea, which possesses only the pain sense. the sensory nerve endings are in the form of branches of axis cylinders among the epithelial cells. Since by a section of the skin parallel with the surface the touch spots can be removed, leaving the pain spots, Waterston concludes that the sensory end organs for pain are more deeply placed than those for touch, and he identifies pain with Meissner corpuscles. It would be possible, however, to explain Waterston's experiment on the basis of the usual assumptions (*i.e.* that pain originates in free nerve terminals and touch in Meissner corpuscles), by supposing that the sensation of pain which remains is due to irritation of the cut ends of the sensory nerves and not to stimulation of the Meissner corpuscles. Further experiments are required to clear up these points. We have at present no evidence as to the histological character of the organs by which the sensations of heat and cold are aroused. They are believed to be Krause's end-bulbs.

For Reference

HEAD (1920). "Studies in Neurology." 2 Vols. (London: Oxford Medical Publishers.)

WALSHE (1942). The Anatomy and Physiology of Cutaneous Sensibility. *Brain*, **65**, 48.

WOLFF and HARDY (1947). On the Nature of Pain. *Physiol. Rev.*, **27**, 167.

CHAPTER XXVIII

SENSATIONS OF SMELL AND TASTE

EVERY living organism shows a susceptibility, *i.e.* a power of reaction, to chemical stimuli. These movements of attraction and repulsion are spoken of as positive and negative chemiotaxis respectively. The aggregation of leucocytes round microbes or other foreign particles in the tissues is also determined by their chemiotactic sensibility. *Chemiotaxis*, then, represents the faculty by means of which these minute organisms are able to adapt themselves to chemical changes in their environment and to react to chemical substances diffusing towards them from a distance.

On the other hand, a chemical sensibility of the body surface, or part of it, furnishes the criterion by which particles are accepted and ingested as food or rejected as useless or harmful. Consciousness in this case would be of something in contact with and affecting some part of the organism itself. The sensation would not be projected farther than the periphery of the body.

These two kinds of chemical sense—the projected and the surface sense—are found throughout almost all classes of the animal kingdom, and in the higher animals at least are known as the senses of smell and taste. The former sense in many animals attains a high degree of complexity and is prepotent in determining the behaviour of an animal in response to the changes in its surroundings. In the elasmobranch fishes the olfactory lobes form the greater part of the higher brain, and extirpation of them produces a loss of spontaneity and of delayed reactions similar to that which can be brought about in higher types by extirpation of the whole of the cerebral hemispheres.

The sense of taste, on the other hand, is used only for sampling the nature of substances taken into the mouth and determining their ingestion or rejection. It is therefore much simpler in its extent and more susceptible of analysis.

THE SENSE OF TASTE

The end organs which subserve the function of taste are represented by the *taste buds*. These are oval bodies (Fig. 319) embedded in the stratified epithelium, which occur scattered over the tongue, a few being also found on the hard palate, the anterior pillars of the fauces, the tonsils, the back of the pharynx, the larynx, and the inner surface of the cheek. On the tongue they are found chiefly in the grooves around the circumvallate papillæ of man, and in the grooves of the papillæ foliatæ of rabbits. A few are also present in between the fungiform papillæ. They consist of medullary and cortical parts, the latter being composed of columnar or sustentacular cells, the former of thin fusiform cells, the taste cells proper. The nerve fibres concerned with taste end in arborisations among these taste cells. The peripheral end of the fusiform cell projects as a delicate process through the orifice of the taste bud, so that it can come in contact with the fluids contained in the cavity of the mouth. A sapid substance, to stimulate these organs, must be in solution ; hence quinine in powder is almost tasteless, owing to its slight solubility.

DIFFERENTIATION OF TASTE. The number of classes of tastes is very limited. We distinguish four primitive taste sensations, viz. sweet, sour, bitter and salt, some authors adding to these an alkaline taste and a metallic taste. Many substances owe their distinctive character when taken into the mouth to the fact that they stimulate not only the taste nerves but also the nerve endings of tactile sensation. Thus acids, when in weak solution, have an astringent character besides their sour taste, and if strong produce a burning sensation. The primitive taste sensations can affect one another if excited simultaneously. With weak stimulation one taste may practically annul another. Thus a dilute solution of sugar is rendered almost tasteless by the addition to it of a few grains of common salt. If the primitive taste sensations are more strongly excited we get a mixed sensation, in which the components can still be distinguished. Thus, adding sugar to lemon juice not only diminishes its acidity but produces a mixed sensation, the quality of which is pleasant and in which the components, sour and sweet, can be easily distinguished. We get no such fusing of sensations as in the eye, where a sensation of white light may result from stimulation of the retina by two complementary colours. Stimulation of one kind of taste organ heightens the sensibility of the other taste organs. Thus, after the application of salt, distilled water may taste sweet. The sense of taste may be investigated in three ways: (1) By ascertaining the greatest dilution of a substance that will cause a definite taste; (2) by ascertaining how long a solution of a substance must be applied for a taste to be aroused; (3) by ascertaining the least difference between the strengths of two solutions of the same substance which taste different to the experimenter.

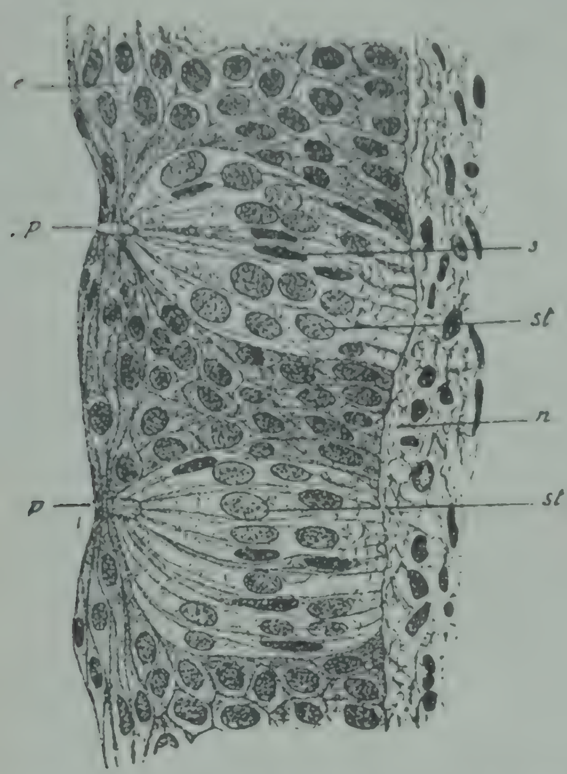


FIG. 319. Two Taste Buds from the Tongue. (KÖLLIKER.)

e. Stratified epithelium.
p. Opening or pore of taste bud.
n. Sub-epithelial tissue.
s. Gustatory cells.
st. Sustentacular cells.

By these methods we find:—

(*a*) The tongue is not equally sensitive at all points to all four tastes. Thus the back of the tongue is more sensitive to bitter, while the tip and sides of the tongue react more easily to sweet and sour substances. A difference may even be detected between the circumvallate papillæ themselves; a mixture of quinine and sugar applied to one papilla may excite chiefly a bitter taste, while with an adjacent papilla a sweet taste may predominate.

(*b*) By certain drugs we can depress the sensibility of the taste organs, and we then find that the various tastes are affected to different degrees. Thus on painting the tongue with cocaine the first effect is a diminution of tactile and pain sensibility, so that the application of acid evokes a very sour taste without any of the astringent or stinging sensations normally aroused by the contact with the acid. After this point the taste sensations are also mostly abolished. The bitter sensation disappears first, then the sweet and then the sour, while the taste of salt appears to remain unaffected. On the other hand, if the leaves of *Gymnema sylvestre* be chewed, the sensations of bitter and sweet are abolished, leaving intact the acid and salt tastes, and also the general sensibility of the mucous membrane.

TASTE AND CHEMICAL CONSTITUTION. There is no doubt that the stimulating effect of any chemical substance on the taste nerves has relation to its chemical constitution. Thus a sour taste is determined by the presence of H ions; the alkaline taste by that of OH ions. The fact that certain acids, *e.g.* acetic, have a stronger sour taste than would correspond to their dissociation, *i.e.* to the number of H ions present, is due to the fact that these acids penetrate more easily into the gustatory cells than the mineral acids with a larger dissociation co-efficient. All the α-amino-acids have a sweet taste. On the other hand, the polypeptides produced by the combination of these amino-acids, as well as the peptones derived from the hydrolysis of proteins, have a bitter taste. Most of the alcohols and sugars have a sweet taste, while the metallic derivatives of these substances are bitter. We do not yet understand the law which determines whether any given substance shall have a taste at all, and what its taste should be.

The nerves of taste are the glossopharyngeal, which supplies the back part of the tongue, and the lingual branch of the fifth nerve and the chorda tympani which supply the front part. All these fibres are probably connected with a continuous column of grey matter in the brain stem, which represents the splanchnic afferent nucleus of the fifth nerve, the nervus intermedius and the glossopharyngeal. Some authors have stated that all the taste fibres of the fifth nerve are derived from the glossopharyngeal by the communication through the tympanic plexus and the chorda tympani nerve, while Gowers has recorded a case of complete unilateral loss of taste in which there was a lesion destroying the fifth nerve, the glossopharyngeal being intact. It seems possible that the actual region of the taste nerve may vary, the fibres running to the splanchnic column of grey matter being contained sometimes in the fifth, sometimes in the glossopharyngeal, and sometimes in both.

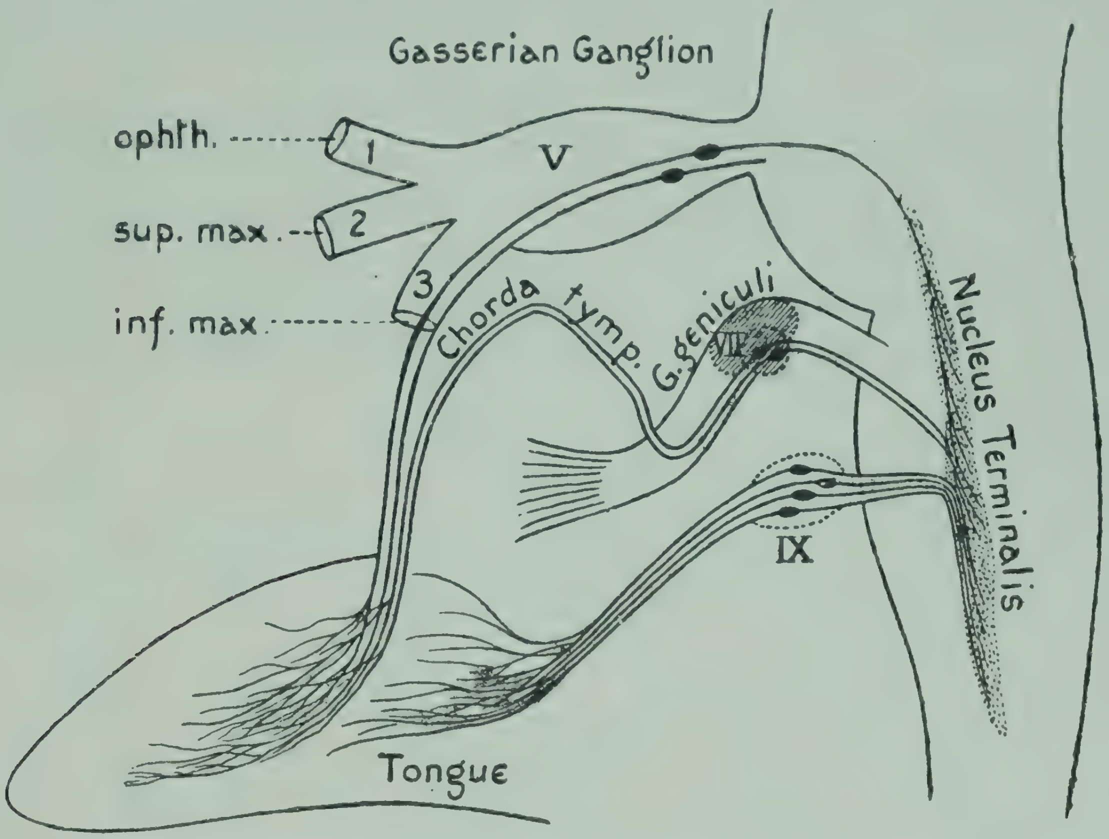


FIG. 320. Diagram showing Origin and Course of the Nerve Fibres of Taste.

Most of our so-called tastes should rather be designated flavours and are dependent not on the gustatory nerves but on the sense of smell. When the olfactory sense is destroyed, very little difference is to be perceived between an onion and an apple. The epicure with a fine palate has really educated his sense of smell and would be but little satisfied with the simple sensations derived from his tongue.

THE SENSE OF SMELL

The analysis of olfactory sensations is rendered difficult by the fact that this sense in man plays but a small part in his usual adaptations. We have thus to deal with a sense which is in many respects vestigial. We see traces of great complexity in its possibilities of performance, but are baffled in our endeavours to reduce the whole of the phenomena to the simpler factors of which they are composed. Moreover, like all vestigial functions, the extent to which the sense is developed varies from one individual to another. Many, for instance, are unable to appreciate the smell of vanilla, of hydrocyanic acid or of violets. On the other hand, in animals such as the dog, the olfactory sense seems to play a great part in determining behaviour, and the nervous associations, which are the physiological basis of ideas, must in these animals be largely connected with olfactory impressions. Another factor which diminishes the importance of olfactory sensations in man is the ease with which the sense organ becomes fatigued. It often happens that the inmates of a room are perfectly comfortable and may perceive no fault in the ventilation, although a newcomer from the outside at once remarks that the air is foul.

THE ORGAN OF SMELL is situated at the upper part of the nasal cavities. Here the mucous membrane covering the superior and middle turbinate bones and the corresponding part of the septum is different from that covering the rest of the nasal passages. Over the lower parts of the nasal cavities the mucous membrane is of the ordinary respiratory type, and is composed of ciliated columnar epithelium containing a number of goblet cells. In the olfactory part the epithelium is much thicker, of a yellow colour, and apparently composed of a layer of columnar cells resting on several layers of nuclei. These nuclei belong to the olfactory cells proper, true spindle-shaped nerve cells with one process extending towards the mucus covering the free surface, while the other is continued along channels in the bone, and through the cribriform plate as one of the non-medullated olfactory nerve fibres. These nerve fibres enter the olfactory lobes, where they terminate by a much-branched arborisation or end basket in the so-called olfactory glomeruli, in close connection with a similarly branched dendrite of the large *mitral* cells of the olfactory lobe. The axons from these latter carry the olfactory impulse towards the rest of the brain. In the connective tissue basis (dermis) of the mucous membrane are a number of small serous glands (Bowman's glands) whose office it is to keep the surface of the membrane constantly moist.

In ordinary respiration the stream of air never passes higher than the anterior inferior border of the superior turbinate bone, so that it does not come in contact with the olfactory mucous membrane. The sensations of smell which are aroused during ordinary respiration depend on diffusion from the respiratory air into the still air of the upper olfactory portion of the nasal cavity. The direction of olfactory attention is achieved by sniffing; in this act the nostrils are dilated and the direction of the anterior part of the nasal respiratory chamber altered, so that the stream of entering air is directed towards the upper olfactory portion of the cavity.

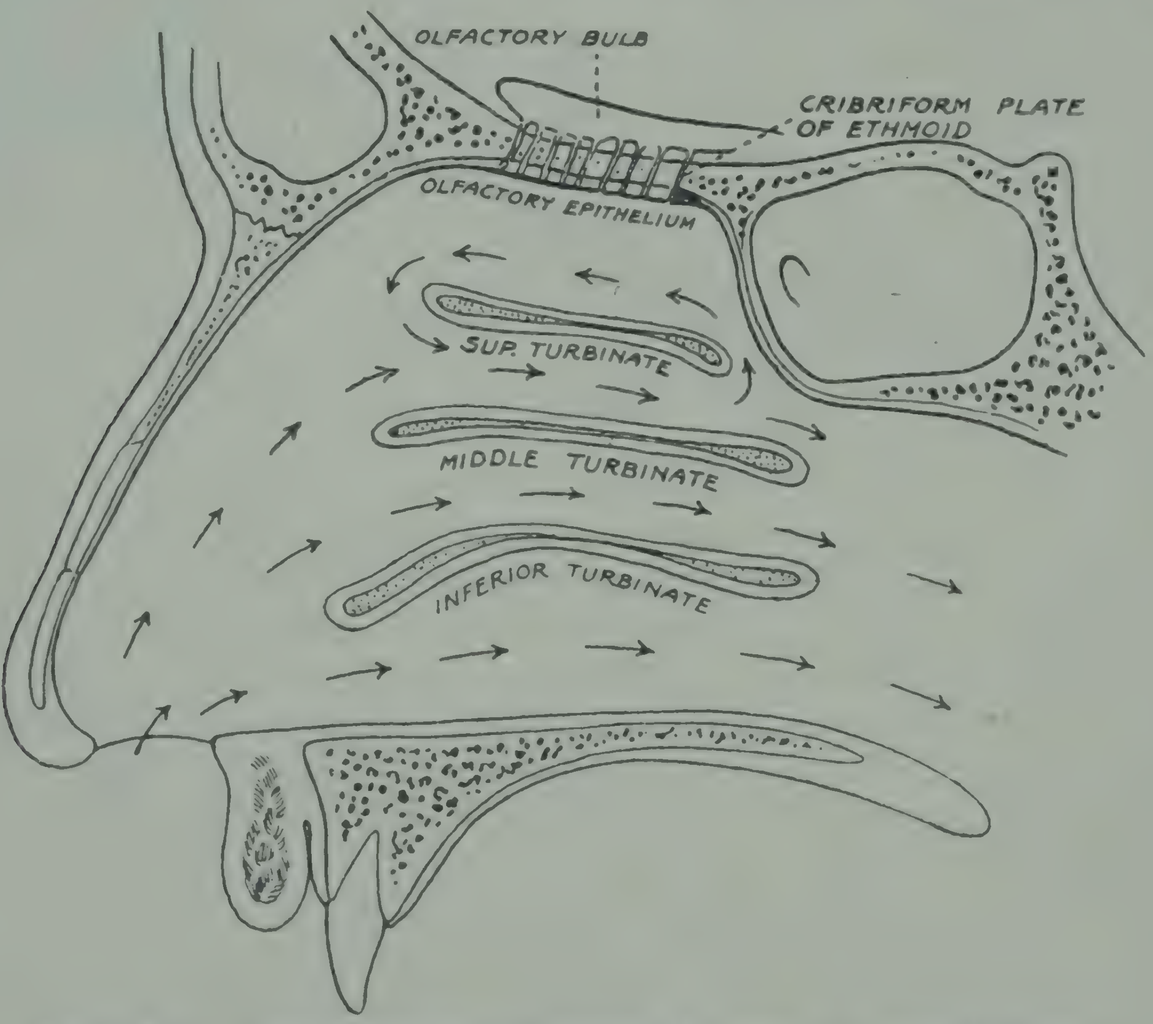


FIG. 321. Antero-posterior Section through the Nasal Fossæ.
The arrows show the direction of the air currents during inspiration.

The fact that the air, which enters the nasal cavity during respiration, does not come into direct relationship with the olfactory epithelium has the following advantages :—

(1) The cold inspired air does not come into contact with and cause damage to the sensory surface.

(2) Foreign particles carried by the air (including bacteria) do not get deposited there. The position of the epithelium at the very top of the nasal cavity is an additional safeguard.

(3) The olfactory epithelium is not dried by the rush of dry air across it.

(4) Noxious vapours only reach it indirectly and therefore do not cause permanent damage as they otherwise might.

The amount of substance necessary to excite sensation is extremely minute. Thus 0·01 mg. of mercaptan diffused in 230 cubic metres of air is still distinctly perceptible. In this case a litre of air would contain only ·00000004 mg. of the substance, and the amount actually in contact with the olfactory epithelium would be still smaller. It is possible, however, to show the presence of these odorous substances in air by physical means. Tyndall pointed out that air containing a small proportion of odorous substances absorbed radiant heat to a much greater degree than did pure air. Thus in one experiment, air containing patchouli absorbed radiant heat thirty-two times as strongly as the pure air. Most odorous substances possess large molecules and have therefore high vapour densities. On this account the smell tends to hang about objects, the rate of diffusion of the vapour being only small.

MODE OF ACTION OF ODOURS. Since the endings of the olfactory cells are bathed in fluid, it is evident that the odorous substances must be dissolved by this fluid before they can excite the olfactory nerve fibres, and in the case of aquatic animals we know that the projected chemical sense, which we call smell, can be aroused only by substances in solution. It is difficult to show in man that the nerve endings can be excited by solutions. Most of the experiments have been made with solutions which had an injurious effect upon the olfactory epithelium. According to Aronsohn it is possible to excite sensations of smell if the nasal cavity be filled with normal saline fluid containing a very small proportion of the odorous substance.

Many olfactory stimuli owe their peculiar character to the simultaneous stimulation of other kinds of nerve endings. Thus a pungent smell, as that of ammonia, chlorine, etc., besides stimulation of the olfactory nerve, involves stimulation of the nerve endings of tactile sensibility, supplied by the fifth nerve.

OLFACTOMETRY. Elsburg, Levy and Brewer have recently devised a satisfactory method of measuring the acuity of smell in different people. A bottle containing the odorous substance can be filled with air at a slightly raised pressure so that on releasing a valve a known volume of air containing the odour is injected into the nasal cavities. For different odorous substances, different volumes of air must be used to reach the threshold, as the following table shows :—

Benzene	5·26 c.c.
Xylene	10·0 ,,
Camphor	15·0 ,,
Clove oil	17·22 ,,
Cedar wood oil	32·65 ,,

They have found that in women just prior to menstruation the acuity of the sense of smell is increased. A cold in the head depresses it often for several days. Smelling a substance depresses for a time the acuity for that substance.

THE CLASSIFICATION OF SMELLS. No satisfactory classification of smells has yet been made. The following facts tend to show that there are a number of primitive sensations of smell, as of other sensations :

(*a*) Certain individuals, whose olfactory sense is in other respects normal, have no power of distinguishing some odours.

(*b*) The olfactory sense is easily fatigued. If it be fatigued so as to be absolutely insensitive for one kind of smell, it is still normally excitable for other smells.

(*c*) It is possible by mixing odoriferous substances in certain proportions to annul their effect on the olfactory organ. Thus 4 grm. of iodoform in 200 grm. of Peruvian balsam is almost odourless, and the same neutralisation of odours is obtained if the odour of each substance be allowed to act separately on each side by tubes inserted into each nostril.

For Reference

PARKER (1922). "Smell, Taste and Allied Senses in the Vertebrates." (Philadelphia and London : Lippincott Co.)

BOOK V

SYSTEMS FOR DISTRIBUTION OF MATERIALS

BLOOD, CIRCULATION AND RESPIRATION

HISTORICAL NOTE

Blood. Although the more conspicuous properties of the blood must have been familiar from the earliest times, it was regarded as a homogeneous fluid until late in the 17th century, when the red cells were first accurately described by Leeuwenhoek (1674). Their enumeration was first effected by Vierordt in 1852, by counting the cells in a film made with a minute known volume of blood: the counting chamber was invented by Cramer in 1855, and the first hæmocytometer somewhat on present lines was that of Gowers (1877) and the now familiar dilution pipette was devised by Potain in 1873.

The arresting phenomena of blood coagulation were prematurely studied by Galen; a speculation by Borelli that it was due to the throwing down of a solid from the plasma was confirmed by Hewson in 1772, and by Buchanan (1845), who showed that fibrin was formed by the action of a rennet-like substance on a soluble precursor. This precursor, fibrinogen, was isolated by Denis in 1856, and a "fibrin ferment" recognised by Schmidt in 1872; Hammarsten, in 1876, showed that fibrinogen was the sole precursor of fibrin, the thromboplastic substance making no contribution. The necessity of calcium salts for clotting was recognised by Arthus and Pagès in 1890, and the existence of prothrombin and its conversion to thrombin by Ca salts by Pekelharing in 1891. The part played by the thromboplastic substance emerged from the work of Wooldridge (1891) and of Morawitz (1903): the latter called it thrombokinase, and put forward a theory of coagulation which is still current. The inhibitory influence of the endothelial lining of the blood vessels was observed by Brücke (1857), the liver regarded as the site of formation of an anti-coagulative substance by Hédon and Delezenne in 1892, and heparin discovered in 1928 by Howell.

Attempts to ascertain the volume of blood in the body have been made since the times of Harvey and of Hales, but Allen Moulin in 1654 and Floyer in 1710 got somewhere near the truth, with figures of one-twentieth and one-fourteenth of the body weight respectively. Experiments on man began, in 1855, with Bischoff's application, to executed criminals, of Welker's rinsing-out method. In 1838, Valentin conceived the principle of adding a given amount of a recognisable substance to the circulating blood, and then estimating its amount there: the first suitable substance to be found, however, was carbon monoxide, used by Gréhant and Quinquaud in 1882.

Blood transfusion, now so popular, was speculated upon by the ancients, and often essayed in the past, *e.g.* by Lower, 1666, but was generally discredited as being inapplicable to man until after the principle of compatibility was revealed by Landsteiner's discovery of the blood groups in 1901.

Haller (1754) mentions that the insoluble part of blood ash contains some particles which are attracted by the lodestone, but the significance of this fact was not realised for a long time. Hæmoglobin was first properly described by Hoppe-Seyler between 1866 and 1871, crystals of hæmoglobin were prepared quite early, notably by Preyer in 1871, and the spectra of the important blood pigments were further studied by Stokes (1864) and Hüfner (1879). Claude Bernard had shown as early as 1854 that CO was poisonous because it was fixed by, and drove out oxygen from, the blood, and Donders in 1872 showed that the combination was a reversible one due to the formation of carboxyhæmoglobin.

Lymph. The lacteal vessels were discovered by Aselli in 1662, and considered by him to play a part in absorption, taking the absorbed material, in accordance with the Galenical doctrine, to the liver. Accordingly, it was not until after the discovery of the thoracic duct by Pecquet in 1651, and of the general lymphatic system by Bartholin (1653) that the system was considered as a whole. Most of our modern views on lymph formation, and on its relation to the circulatory system, originated in Starling's work about 1894.

Circulation. For five centuries from the time of ERASISTRATUS of Alexandria (3rd century B.C.), the arteries were believed to be filled with air, and only the veins with blood. In the 2nd century A.D., GALEN showed by experiment that the arteries were quite filled with blood: from his further researches was formulated his doctrine of *pneuma*, or spirits, which was accepted with little modification during the next 14 centuries, and which was briefly as follows: food first becomes "chyle" in the alimentary canal, is absorbed into the blood, and, being carried at once to the liver, is there converted into blood and provided with "natural spirits"—a nutritive principle. The blood, thus endowed, and mixed with blood of the vena cava, is distributed all over the body by way of the veins, in an ebb-and-flow motion, and also reaches and recedes from the right ventricle. A part of it, however, passes from the right ventricle by invisible pores into the left side, and another part passes to and fro along the pulmonary artery or "arterial vein," and, reaching the lungs, which it nourishes, at the same time gives out into the expired air all its noxious ("fuliginous") vapours. That part of the blood from the right side which seeped through into the left side is next mixed with air, which, drawn into the lungs in inspiration, and passing along the pulmonary vein, or "venal artery," reaches the left side of the heart, where heat is created, and the "vital spirits" engendered, giving to the blood its bright colour. The vital spirits were necessary for life and movement, and the vitalised blood ebbed and flowed along the arteries with their pulsations, expansion (diastole) of the heart and arteries occurring together at one phase, and collapse (systole) at another. When the blood reached the brain, the "animal spirits" were generated: thus pneuma, rid of air and blood, could pass down the nerves to cause movement and sensation.

One of the chief points in the Galenical doctrine concerned the route by which the blood of the right side of the heart reached the left side. Doubt was first thrown on the permeability of the interventricular septum by VESALIUS in 1543, and was confirmed by REALDUS COLUMBUS in 1559. Next, the disposition and action of the valves of the veins were described by FABRICIUS at Padua in 1574: the logical inference from his discovery was, however, only pointed out later by his pupil, WILLIAM HARVEY.

The fact that blood can pass from the right to the left ventricles by way of the lungs—the pulmonary circulation—was first advanced by IBN NAFIS, an Arabian physician, in the 12th century,* and again independently by MICHAEL SERVETUS in 1553,† and by COLUMBUS and three others in the 16th century. SERVETUS, however, only believed that part of the blood passed through the lungs.‡

The discovery that the blood travels round in a circuit, returning again and again to the same point, was made by WILLIAM HARVEY, physician to Charles I., and was given to the world in a book of seventy-two pages, printed at Frankfurt in 1628.§

This, the greatest discovery in the whole of medical science, of which indeed it laid the foundation, was the result of prolonged experiment and patient observation; in his masterly though modest treatise Harvey shows that the circulation is a logical necessity to which we must be led by the contemplation of certain elementary facts, *e.g.*:—

(i) The arrangement of the valves in the heart and veins is such as to allow of one-way movement only.

(ii) "If the heart is grasped in the hand, it may be felt to become harder during its action." When this happens, the ventricles become smaller and the arteries *expand*; if an artery is severed, blood spurts from its central end at each heart contraction.

(iii) Since, at each contraction of the heart, blood is forced into the arteries and cannot return, it follows that, even if only a small quantity were so expelled, there would not be enough blood in the heart, or even in the whole body, for the process to continue, nor could all this blood remain in the arteries. The heart must, therefore, receive blood continuously from the veins and pass it out continuously into the arteries. If a vein is compressed or tied by a ligature, it swells up on the side furthest from the heart, and collapses on the side nearest the heart. With an artery it is the reverse. Hence the blood must be passing continuously from the arteries, through the tissues into the veins. Similarly, the blood in the right ventricle must pass by way of the pulmonary artery and lungs to the pulmonary veins and so reach the left side of the heart.

Harvey could not be certain how the blood passed through the tissues from arteries to veins, but with the use of the microscope, the capillaries were discovered by Malpighi in 1661, and the missing link in the argument was thus provided.

Once the fact of the circulation was established, further aspects soon pressed for attention. LOWER (1669) made important observations on the structure and action of the heart, and BORELLI (1680) considered the circulation from the hydraulic standpoint: but the next important experimental study was made by the Rev. STEPHEN HALES (1733). By inserting vertical tubes into the arteries or veins in horses and dogs, he showed the relative pressures in those vessels, distinguished end-on from lateral pressures, and, in confirmation of FLOYER (1710), explained the loss of pulse in the veins. Hales also tried to assess the cardiac output, carried out perfusions, estimated the velocity of the blood, and made studies on œdema production Attempts to understand the significance of the heart sounds began with LAËNNEC, who invented the stethoscope, in 1819. That the apex of the heart moved but little during the beat was a forgotten observation of LEONARDO DA VINCI in the 15th century, rediscovered independently by HAYCRAFT in 1891 by an almost identical method.

---

* SAMI HADDAD and AMIN KHAIRALLAH (1936). *Annals of Surgery*, **104**, 1; MAX MEYERHOF (1936). *Isis*, **23**, 100. (1933). *Bull. de l'instit. d'Égypte*, **16**, 33.

† OSLER (1909). "Michael Servetus". (London, Oxford University Press.)

‡ IZQUIERDO. *Bull. Inst. Hist. Med.*, 1937, **5**, 914.

§ HARVEY. "Exercitatio Anatomica de motu cordis et sanguinis in animalibus," Frankfurt, 1628. Many editions of English translations may be had.

D'ARCY POWER (1897). "William Harvey." (London, Unwin.)

SINGER (1922). "The Discovery of the Circulation of the Blood." (London, Bell.)

The volume of the cardiac output was for centuries a matter for speculation : a method for its determination was outlined by FICK in 1870, but was not much used until 1912 when KROGH and LINDHARD introduced the nitrous oxide method.

The frog's heart, which was used by Harvey, and again in 1797 by v. HUMBOLDT, has played a large part in the investigation of the properties of heart muscle. The myogenic view of the origin of the heart beat, which had been held by HALLER (1736) was challenged by LEGALLOIS (1812), who revived the idea of a neurogenic explanation : in apparent support of this, various local ganglia were discovered by REMAK (1844), LUDWIG (1848) and BIDDER (1852), and these came, through Bidder and ECKHARD (1858) to be regarded as local motor centres : the experiments of STANNIUS (1852), though now differently interpreted, are still well known. A fierce and prolonged controversy raged, but was finally settled in favour of the myogenic theory, largely by the work of GASKELL (1882–7). The all-or-none property was discovered by BOWDITCH in 1871, and the refractory phase by KRONECKER and STIRLING in 1874 ; the influence of ions by RINGER in 1880, the nature of the conducting system (A–V bundle) by KENT (1892) and the AV node by KEITH and FLACK in 1906. The electrical changes in the human heart were first recorded by A. D. WALLER in 1887, with the capillary electrometer. The string galvanometer, invented in 1901, was used for recording the electrocardiogram by EINTHOVEN in 1903.

Although LEGALLOIS (1812) and PHILIP (1826) showed that stimulation of the spinal cord accelerated, and of the medulla slowed, the heart rate, the pathways were unknown till 1845, when the WEBER brothers showed the inhibitory nature of the vagus : the effect of the sympathetic was shown by v. BEZOLD in 1862, and the action of the depressor nerve by CYON and LUDWIG in 1866. The importance of the carotid sinus was not revealed until 1925 (KOCH).

A method for the study of the mammalian heart doing normal and controllable amounts of work, while fed with blood, yet freed from the complications of the variable factors introduced by the systemic circulation, was long sought and often attempted, but for long with only partial success. The practical problem was solved by the work of KNOWLTON and STARLING in 1912, in which the heart-lung preparation was introduced ; this has proved of value for a variety of investigations.

The mercurial manometer was introduced by POISEUILLE (1825) and the float for recording by LUDWIG in 1847, and thereafter studies on arterial pressure quickly developed. LUDWIG also introduced the kymograph, the stromuhr (1867) and other experimental devices. The study of intracardiac pressures by CHAUVEAU and MAREY in 1861 led to a more accurate analysis of the events in the cardiac cycle, while FRANK's work, about 1906, on the physical properties and defects of recording apparatus marked a new epoch in the technique of graphic registration. The apparatus for measuring human arterial pressure was introduced by RIVA-ROCCI in 1896. The dynamics of the flow of blood began to be studied by POISEUILLE in 1840, and between 1825 and 1875 a good deal of attention was devoted to the study of the form and velocity of the arterial pulse by the WEBER brothers, DONDERS, and MAREY, the first sphygmograph being produced by the last-named in 1860. The venous pulse was known from ancient times, but was not adequately investigated until MACKENZIE's studies in 1902.

The variability of the peripheral resistance through contraction of the small arteries was not properly appreciated until the 19th century, but was firmly established when the hitherto hypothetical vasomotor nerves were found in 1851 by CLAUDE BERNARD and others. That the capillaries were also independently contractile was claimed by ROUGET in 1879, but was not recognised generally until the work of KROGH in 1916, and the subsequent work on human skin by LEWIS.

Respiration. Aristotle, circa 300 B.C., thought that respiration cooled the blood, and ERASISTRATUS, who postulated two pneumata, thought that one of these, the vital spirit, was a form of air which filled the arteries, and was derived from air entering the left heart from the lungs. GALEN, about A.D. 170, showed that arteries were, in fact, entirely filled with blood, and was the first to suggest that breathing not only added something to the blood (vital spirits), but also removed something (" fuliginous vapours ") from it at expiration. He recognised the chief respiratory muscles, and knew the functions of the phrenic and recurrent laryngeal nerves. LEONARDO DA VINCI, about A.D. 1500, recognised the similarity between flame and living things, in that both consumed and altered the air.

It is evident that, even after the discovery of the circulation of the blood, in 1628, the function of breathing could not in any case have been appreciated until the discovery of oxygen, of carbon dioxide, and of the fundamental nature of combustion. In fact, the meandering trail of discovery took over 160 further years to pass those milestones.

It is true, however, that on the way a few exceptional observers had a misty pre-view of things to come, though this was overlooked. So, BOYLE, in 1670, by the use of the air-pump, demonstrated the indispensability of air for life, as for flame ; he extracted gases from the blood, and claimed that a part of the air was essential for life and combustion. So, also, **Mayow**, in 1668, probably under the influence of Boyle, not only actually discovered what we now know as oxygen (his " nitro-aerial spirit "), but came within an ace of explaining, once and for all, the true nature of combustion, of respiration and of the source of animal heat. The dominance, in chemistry, of the phlogiston theory of Stahl from about 1680 to 1794, however, was an obstacle to the acceptance of such views.

One of Mayow's accepted discoveries (1668) was that of the manner in which the expansion of the chest draws air into the lungs. Opening of the chest was long known to cause death, but **Vesalius**, about 1543, showed that death could be warded off if the lungs were rhythmically inflated by bellows. The regular inflation and deflation of the lungs was afterwards supposed in some way to be essential, until **Hooke**, in 1667, showed that, provided the air was renewed, as by a continuous blast which escaped by holes made in the lungs, the actual movements of the lungs could be dispensed with

That the change from venous to arterial blood took place in the lungs, and was connected with the renewal of the air in them, was discovered by **Lower** in 1669 : he showed that the colour change of the blood could be observed in animals with opened chests, that it took place in passing through the lungs, but ceased soon after artificial respiration was stopped.

The definitive discovery of carbon dioxide was made by **Black** in 1754 (though it was confused with nitrogen until 1772), and oxygen was discovered independently by **Priestley** and by **Scheele** in 1771. Many facts regarding combustion, respiration, and the formation of oxygen by plants were discovered by Priestley, but owing to his adherence to the phlogiston theory, he missed the correct interpretation of them. It was **Lavoisier** who supplied this, in 1775, and between then and 1794, when he was guillotined, he unveiled the true nature of combustion and respiration (the composition of water having been discovered by **Cavendish** in 1781, and of air in 1784), outlined the changes which took place in the air as a result of breathing, determined the oxygen usage of men at rest and doing work, propounded the true cause of animal heat, and, in general, laid the foundations for the study of oxidative metabolism. He thought, however, that the lungs were the actual site of the combustion process.

Edwards (1824), following the work of **Spallanzani** at the close of the 17th century, showed that animals, or their tissues, when deprived of oxygen, continue to give up carbon dioxide, and he formulated a theory of respiratory exchanges much akin to our present-day views. The fuller explanation had to await the further study of the blood gases. These had been extracted *in vacuo* by Boyle, but the first thorough investigation was that of **Magnus** in 1837. Using a mercurial pump, he extracted the gases from blood, demonstrated the difference between arterial and venous blood, and thus established the true relations of the blood, air, and general tissues of the body. The laborious method of extraction *in vacuo* was thereafter employed for the estimation of the blood gases up till about 1900, when the ferricyanide method was introduced by J. S. **Haldane**.

The study of the tensions of the blood gases began with **Lothar Meyer**'s discovery, in 1857, that oxygen was not released from blood in proportion to any given reduction of pressure, but much more slowly, and was, therefore, not in simple solution. **Hoppe-Seyler**, five years later, showed that it was held in combination with hæmoglobin, and in 1885 **Bohr** expressed the relationship between oxygen content and oxygen partial pressure in the form of the oxygen-dissociation curve. The means by which CO_2 was carried, adumbrated by **Zuntz** in 1867, was not explained till **Hamburger**'s work in 1892. The remarkable interrelations between various blood components, which found its final expression in the nomographic work of **Lawrence Henderson** about 1928, began with the observations, by **Bohr**, **Hasselbalch** and **Krogh** in 1904, that carbon dioxide tended to expel oxygen from the blood, and by **Christiansen**, **Douglas** and **Haldane** in 1914, that oxygen reciprocally tended to expel CO_2.

Long controversy centred around the ultimate cause of the respiratory exchange in the lungs, and began with **Müller**'s declaration, in 1870, that the interchange was a secretory process, a position taken up and held consistently by **Haldane** from 1896 to 1922. The invention of the aerotonometer by **Pflüger** in 1869 enabled the matter to be tested by various investigators, and finally, in 1908, by **Krogh**, with his microtonometer, which seemed to decide the issue in favour of a physical explanation. **Haldane** being still recalcitrant, however, J. **Barcroft** entered the lists, and after the contending parties had, in turn, scaled higher and higher peaks in their study of the effects of altitude, and exposed themselves and their colleagues to progressively reduced pressures or lowered oxygen atmospheres in chambers, it came to be almost universally agreed that the oxygen intake is, in fact, effected as a result of a difference in partial pressures, and not as an act of secretion of oxygen from air to blood.

That the movements of respiration are dependent upon a centre in the medulla was first shown by **Legallois** in 1811, and further definition and studies of the respiratory centre, and of the effects of vagus section, were due to **Flourens** (1842), **Gad** (1892), and **Rosenthal** (1864), while **Hering** and **Breuer**, in 1868, demonstrated the effects of inflation and collapse of the lung, and of the influence of vagal section thereon.

The influence of the chemical composition of the blood on the action of the respiratory centre was noticed by **Rosenthal** in 1862, in studies on hyperventilation apnœa : this he attributed to over-oxygenation of the blood, but **Traube**, in 1867, showed that over-ventilation with nitrogen would also cause apnœa, and **Pflüger**, in 1868, showed that increase of CO_2, as well as deficit of oxygen, could excite the centre. The classical work of J. S. Haldane did much to clarify the means by which breathing is regulated, and particularly to show that carbon dioxide was the principal chemical regulator of respiration (Haldane and Priestley, 1905), while response to alteration of oxygen partial pressures was much less delicate. The effects of altered atmospheric pressure, studied by Paul Bert in 1878, also provided much valuable information. The discovery that the pressure and composition of the blood flowing through the carotid sinus region exerts an important effect on the respiratory centre was made by Heymans and his colleagues in 1927–30.

CHAPTER XXIX

THE BLOOD

THE function of the blood is to convey food materials and oxygen to the tissues. From these it receives in exchange their waste products, and carries them to the excretory organs, such as the lungs and kidneys, by which they are eliminated. The composition of the blood must vary from time to time and place to place according to the activity and function of the organ which it is traversing, but the organs are adjusted to respond to very minute changes in the composition of the circulating fluid, and add to or subtract from its constituents according as these are present in deficiency or excess. The changes are therefore kept within narrow limits and we may regard the blood as a fluid of approximately constant composition.

Blood obtained from a mammal is an opaque fluid varying in tint according to the vessel from which it is derived, being scarlet when taken from an artery, purplish in colour when taken from a vein, the difference being determined by the degree of oxygenation of the blood. On shaking venous blood with air, it takes up oxygen and acquires the scarlet colour characteristic of arterial blood. If examined in a thin layer under the microscope, its opacity is seen to be due to the fact that it is not homogeneous, but consists of a number of *corpuscles* of different kinds, suspended in about their own bulk of a light yellow transparent fluid, the *plasma*. It is seen to contain three main classes of corpuscles. Much the most numerous and conspicuous are the *red corpuscles*. Next most conspicuous are the *white corpuscles;* the third type of formed elements are very small and are called the *blood platelets*. The average numbers of these formed elements per cubic millimetre of normal human blood are as follows :

Red corpuscles	5,400,000
White corpuscles	9,000
Platelets	300,000 (variable).

In the whole body there are about 32 billion red corpuscles and about 50,000 million white corpuscles. The total red cell volume of the body is about 2 l., or 30 c.c./kg. body weight.

Normally, the process of clotting ensues shortly after the blood has left the vessels. If blood be received into a beaker it is at first perfectly fluid. After three to eight minutes it begins to be viscous, and if poured out of the beaker leaves an adherent layer on the sides of the vessel. A minute later the whole mass of the blood becomes solid and the beaker can be inverted without spilling its contents. If a section be made of this blood clot, it is found to owe its solidity to a network of fine threads of a protein substance named *fibrin*, which have formed throughout the plasma and enclosed the corpuscles in their meshes (Fig. 322). On leaving the clot for some hours, drops of yellow fluid appear on its surface and run together. The whole clot contracts, and finally there is a dark red clot, floating in a transparent yellowish fluid known as *serum*. If the blood on leaving the vessels be whipped with a bunch of twigs, or stirred with a glass rod, the filaments of fibrin as they are formed are deposited on the stirrer. After three or four minutes the twigs can be withdrawn and the spongy fibrin collected. The blood which is left consists only of the corpuscles *plus* serum, and it will not clot since its fibrin has been removed. It is known as *defibrinated*

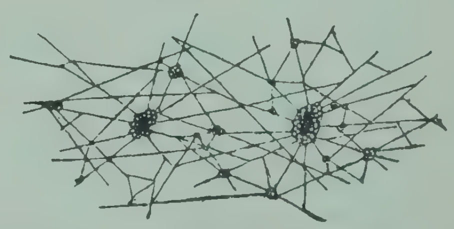


FIG. 322. Network of Fibrin, after washing away the corpuscles from a film of blood that has been allowed to clot; many of the filaments radiate from little clumps of blood platelets. (SCHAFER.)

blood. Since clotting occurs in blood plasma entirely separated from corpuscles, we must look upon the process of coagulation as determined by changes in the blood plasma.

THE QUANTITY AND COMPOSITION OF THE BLOOD IN MAN

The Total Blood Volume. The amount of blood in the body can be estimated by Welker's method, in which the animal is bled and the vessels washed out with salt solution. The depth of colour of blood *plus* washings is then determined and compared with that of the undiluted blood. In this way it was found by Bischoff that in man and the dog the blood equals about 7·7 per cent. of the body weight.

Modern methods, applicable also to man, involve the addition to the circulating blood of a known amount of some substance which does not quickly leave the blood stream, and the subsequent estimation of the substance in the blood after mixing has occurred. Suitable substances are carbon monoxide, an appropriate dye, or a radioactive component of the red cells. The first and third give the volume of the whole blood, the second that of the plasma, from which the volume of blood can be deduced if the hæmatocrit value is known.

The carbon monoxide method,* not now much used, depends on the fact that when carbon monoxide gas is inhaled it combines with hæmoglobin. It therefore depends on the determination of available hæmoglobin which is accessible to the circulation. It probably does not include that present in corpuscles stored in the spleen, though sometimes there may be some increase from this source.† The estimation probably gives too high a result, because some of the CO would be fixed by the myoglobin of the muscles. The error is probably not greater than 10 per cent.

The dye method ‡ consists in injecting a known amount of a non-toxic, non-diffusible dye substance into the blood stream and estimating its dilution with plasma by colorimetric comparison. Many dyes easily diffuse out of the circulation, while colloidal ones are liable to be taken up by the cells of the reticulo-endothelial system. In either case the estimated blood volume would be too high. The dye now used is Evans Blue (T 1824).§ In performing the test a few c.c. of blood are removed from an elbow vein into a syringe containing one-fourth of its volume of isotonic (1·6 per cent.) sodium oxalate to prevent clotting. About 0·2 c.c. per kg. body weight of a 1 per cent. solution of the dye in distilled water is then slowly injected by the same needle. Four minutes later a second specimen of blood is withdrawn into a quarter of its volume of oxalate solution. A part of this sample is drawn into a hæmatocrit tube and centrifuged for twenty minutes at a high speed in order to determine the relative volume of corpuscles and plasma (+ oxalate). From this is calculated the ratio of plasma : plasma + oxalate, since $\frac{1}{5}$th of the whole mixture is oxalate. The percentage by volume of corpuscles and plasma is also calculated. The rest of both samples of blood are centrifuged in order to obtain the plasma. These samples of plasma are then compared by preparing the following mixtures :—

Standard	2 c.c. of a standard dye solution ($\frac{1}{267}$ per cent.).‖ 2 c.c. of the plasma before dye injection. 2 c.c. 0·9 per cent. NaCl solution.
Test	2 c.c. of plasma after dye injection. 4 c.c. 0·9 per cent. NaCl solution.

---

* HALDANE and LORRAIN SMITH. *J. Physiol.*, 1900, **25**, 331. A modified form has been described by CHANG and HARROP, *J. Clin. Investig.*, 1928, **5**, 393.

† WATERFIELD. *J. Physiol.*, 1931, **72**, 117.

‡ KEITH, ROWNTREE and GERAGHTY. *Arch. Int. Med.*, 1915, **16**, 547.

§ DAWSON, EVANS and WHIPPLE. *Amer. J. Physiol.*, 1920, **51**, 232 ; GIBSON and EVANS, *J. Clin. Investig.*, 1937, **16**, 301. CROOKE and MORRIS. *J. Physiol.*, 1942, **101**, 217.

‖ There is no esoteric significance in the curious fraction $\frac{1}{267}$: it merely happens to be a *convenient* strength, obtained by diluting 0·75 c.c. to 200 c.c., as originally recommended.

The two solutions are compared in a colorimeter (preferably photoelectric *) and the strength of the test solution read off as a percentage of the standard. Let x be this percentage, and let V be the c.c. of 1 per cent. dye injected, and C the ratio, c.c. plasma/c.c., plasma + oxalate.

We actually compare the concentration of dye in the plasma (diluted with oxalate), with that of the known dilution of 267 times. For the test to be x per cent. of the standard, the V c.c. of dye must have been diluted, first to the fraction C, and then $\frac{267 \times 100}{x}$ times. Hence the volume of plasma in the circulation must be $\frac{26{,}700\,\mathrm{VC}}{x}$.

The blood volume is calculated from the hæmatocrit reading.

$$\text{Total blood volume} = \frac{100 \times \text{c.c. plasma}}{\text{percentage plasma in blood}}.$$

By reversing the process and injecting undyed plasma into the dyed circulation, errors due to dye losses are claimed to be reduced. This method gives plasma volumes only 85 per cent. of those by dye-distribution. Further, the total cell volume is claimed to be only 85 per cent. of that derived from the hæmatocrit of arterial blood, partly because distribution of cells is not uniform throughout the circulation.†

Radioactive elements also have been used for measurement of blood volume. Some blood is mixed with sodium phosphate containing radioactive P, which the corpuscles absorb. These are well washed and their radioactivity measured : a known amount is injected and the radioactivity of the circulating blood determined after a suitable mixing interval.‡

Estimates of blood volume vary from $\frac{1}{10}$th to $\frac{1}{20}$th of the body weight. The CO method gives 5 to 7 per cent. of body weight, the dye method in most laboratory mammals gives about 7 per cent. of body weight, but no constant relationship to body surface.§ As a round figure we may take 5·0 l. as the blood volume of a 70 kg. adult. In children it is also in linear relation to body weight.||

The total blood volume certainly varies slightly with alterations in the condition of the animal, *e.g.* in thirst. It is influenced by posture, there being a loss of up to 15 per cent. of the total plasma in the erect position as compared with the horizontal.¶ It is also affected by the height of the blood pressure as well as by the oxygen tension in the air breathed, and therefore alters with the altitude. Some of these variations we shall have to consider more fully in a later section. Any lowering of blood pressure causes an absorption of fluid from the tissues into the blood, so that the latter becomes more dilute. The blood during the last stages of slow bleeding may contain many fewer corpuscles than were present in the first samples, pointing to a corresponding dilution of the blood during the bleeding by means of tissue fluid. By this means, *i.e.* the absorption of fluid from tissues, the volume of circulating blood after a limited hæmorrhage is brought up to normal, though in conscious humans the hæmodilution may take as long as 90 hrs. to complete.**

For Reference

ERLANGER (1921). Blood Volume and its Regulation. *Physiol. Rev.*, **1**, 177.

PETERS and VAN SLYKE (1932). Quantitative Clinical Chemistry. (Baillière, Tindall and Cox, London.)

ROWNTREE, BROWN and ROTH (1929). The Volume of the Blood and Plasma. (Saunders, Philadelphia and London.)

---

* KENNEDY and MILLIKAN. *J. Physiol.*, 1938, **92**, 276.

† LAWSON, *et al.* *Am. J. Physiol.*, 1947, **151**, 282–310.

‡ REEVE and VEALL. *J. Physiol.*, 1949, **108**, 12.

§ COURTICE. *J. Physiol.*, 1943, **102**, 290 ; COURTICE and GUNTON. *J. Physiol.*, 1949, **108**, 142.

|| MORSE, *et al.* *Amer. J. Physiol.*, 1947, **151**, 448.

¶ WATERFIELD. *J. Physiol.*, 1931, **72**, 110.

** WALLACE and SHARPEY-SCHAFER. *Lancet*, 1941, ii, 393.

Determination of the Relative Volumes of Corpuscles and Plasma. The relative volumes of corpuscles and plasma in any sample of blood may be approximately determined by means of the *hæmatocrit* (Fig. 323).* This is a graduated capillary tube of uniform bore in which the specimen of blood, treated with a suitable anti-coagulant, is centrifuged at a high speed until the sedimented corpuscular layer shows no further shrinkage, and appears translucent. The depth of the layer of corpuscles is then read off, multiplied by 0·95 to allow for trapped plasma,† and expressed as a percentage of the whole length of the column of blood. Other methods (*e.g.* by adding the blue dye T.1824 in known amount and then finding its concentration in the plasma ; or by diluting the plasma and finding the change in its protein content) give a cell volume about 8·5 per cent. lower than that given by the hæmatocrit.‡ It is usually found, by the hæmatocrit, that in normal human blood the corpuscles occupy about 40 to 50 per cent. of the volume of the blood.

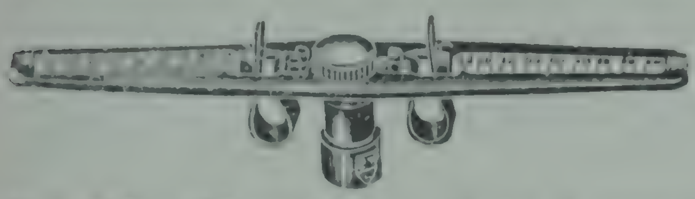


FIG. 323. Hæmatocrit. (From HARRIS' "Experimental Physiology.")

The number of red cells in unit volume of the circulating blood is variable in two ways, viz. by alterations in the quantity of plasma in circulation, as mentioned above, and by variation in the total number of red cells in circulation. The latter may alter abruptly in consequence of the storage in, or liberation of red cells from, the spleen. The significance of the spleen in this respect in man is not known.§

Cell Count and Hæmoglobin Content. In order to enumerate the red corpuscles, a known volume of the blood is diluted with a known amount of an isotonic fluid and the number in a measured volume of the mixture is counted with a hæmocytometer. The average number (*v.* Fig. 324) of red corpuscles is about 5,400,000 per cubic millimetre in healthy adult men, and rather fewer and more variable, in adult women. The enumeration of corpuscles is subject to considerable errors, rarely less than 5 per cent. Moreover, different conditions of the circulation may cause variations in the relative distribution of plasma and corpuscles in different parts of the circulation, so that the blood-count of a specimen from the capillaries of the finger or lobe of the ear may differ considerably from a count in blood obtained directly from a vein or artery.

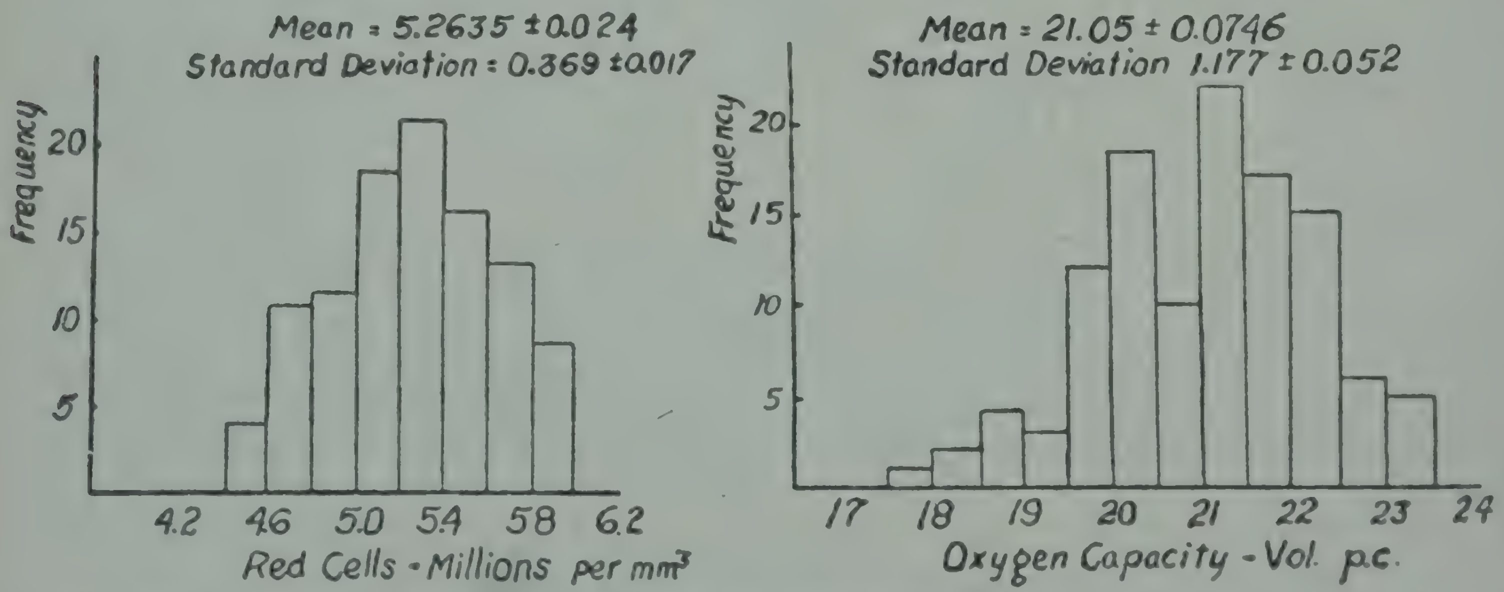


FIG. 324. Showing statistical grouping of red cell count and oxygen capacity in American male students, age 18–30. (FOSTER and RAYMOND JOHNSON.)

More important, perhaps, is the determination of the hæmoglobin. For this purpose, a measured quantity of the blood, say, 20 c.mm., is mixed with a given volume of water. The red fluid thus obtained is compared (often photoelectrically) in a colorimeter, or a specially designed "hæmoglobinometer," with a standard. Usually the hæmoglobin is converted into a more deeply coloured and more stable derivative, *e.g.* cyanhæmoglobin, or acid hæmatin.[*] Haldane's method [†] depends upon the conversion of the hæmoglobin into CO-hæmoglobin, and its comparison with a standard specimen of CO-hæmoglobin, which is stable in solution. Haldane's standard for hæmoglobin in human blood is one with an oxygen capacity now accepted at 19·8 volumes per cent., or 14·8 g. Hb or 49·5 mg. Fe per 100 ml.[‡] This is called 100 per cent. on the Haldane hæmoglobinometer as defined by British Standards Institution. This standard is largely used in clinical work in Great Britain. Normal adult males usually have about 102 to 114 per cent., and females about 88 to 103 on the Haldane scale. The practice of stating the hæmoglobin in terms of g. Hb per 100 c.c. blood, or of oxygen capacity, is increasing.

---

* HEDIN. *Skand. Arch. Physiol.*, 1891, **2,** 134, 360 ; HIROTA. *J. Biophys.*, 1925, **1,** 233.

† BARNES *et al.* *Clin. Sci.*, 1948, **7,** 135.

‡ CHAPIN and ROSS. *Amer. J. Physiol.*, 1942, **137,** 447.

§ BARCROFT *et al.* *J. Physiol.*, 1925, **60,** 443.

Another method is to determine the iron content of the blood,[§] since hæmoglobin contains 0·335 per cent. Fe, and it has been shown that in blood the ratio of oxygen capacity : iron content is 393 ml./g. Fe.

An important clinical determination is that of the *Colour Index* of the blood. For this a red cell count and a hæmoglobin determination are needed. Then, calling 5 million red cells 100 per cent., the colour index $= \dfrac{\text{per cent. Hb}}{\text{per cent. cells}}$ This index, which normally is 1·0, shows the relative hæmoglobin content of the average corpuscle.

The Oxygen Capacity of the Blood. Instead of determining the hæmoglobin, we may measure directly the oxygen capacity of the blood, since the oxygen-binding power of this fluid is entirely dependent on the amount of hæmoglobin it contains. For this purpose we may make use of the fact discovered by Haldane, that the combined oxygen in oxyhæmoglobin is liberated rapidly and completely on addition of a solution of potassium ferricyanide to alkaline laked blood, and may thus be easily measured.

For this purpose the blood-gas apparatus of Haldane is convenient (Fig. 325).

Into the flask B is placed 1·5 ml. of borate buffer of *p*H 10·0.[||] The blood (1 ml.), rendered incoagulable with oxalate, is run into it, and well mixed, laked, and oxygenated by rotating in the flask for a few minutes. Then 0·25 ml. of 6 per cent. ferricyanide solution and a pinch of saponin is put into the small tube C, and the cork inserted, the control flask A, containing 2·5 ml. of water, being similarly corked, and both flasks placed under water in the bath D, the three-way taps J and K being open to the air with the ⊤ bore thus : ⊤. After allowing a few minutes to attain temperature equilibrium, the levelling tubes H and I (filled with water containing a little saponin or bile-salts) are adjusted so as to bring the water in F and G to the marks, and the taps J and K are then turned with the bore thus : ⊣. The burette E is then read, the flask B is tilted so that the ferricyanide runs into the blood solution, and the flask is shaken in the bath until no more gas is evolved. After allowing a few minutes to equalise any differences of temperature, the water level is again adjusted in F and G, and the volume of gas evolved read off on the burette E, and reduced to dry volume at N.T.P. When the *p*H 10 buffer is used the results agree with those obtained by the Van Slyke method.

For most work on blood gases the Van Slyke apparatus [¶] or its modern manometric form [**] is valuable.

The manometric Van Slyke apparatus depends on the principle of measuring the *pressure* of gases evolved or absorbed in a chamber of a *given volume*. From these pressures, or pressure differences, the volumes, or volume differences at N.T.P., can be calculated in accordance with Boyle's law. The *principle* of the apparatus is illustrated by Fig. 326, from which the actual apparatus differs only in some unessential details. The tube A, graduated in mm. is a barometer filled with dry mercury, and connected by a U-shaped bend with the extraction chamber E, junction being effected by a thick-walled rubber tube D. The extraction chamber is graduated at three marks, from the tap, viz. 0·5 ml., 2 ml., and 50 ml., and can be filled or emptied of mercury by raising or lowering the bulb B. If it is filled with mercury, and the tap G closed, lowering of the bulb to a low enough level will produce a vacuum in the chamber. By manipulating the level of the mercury bulb and the tap C, the vacuum can be made of any volume desired, *e.g.*, the mercury can be brought exactly to the 2 ml. mark. If the mercury in the chamber were dry, the levels in the chamber and in the barometer would now be equal; if it was wetted with water this would exert a vapour pressure dependent on the temperature, and the barometer would be higher by that amount, *e.g.* by 17·5 mm. at 20° C. Suppose that we were now to introduce into the chamber (at 0° C.) 0·5 ml. of a gas consisting of, say, 10 per cent. CO_2, 20 per cent. O_2 and 70 per cent. N_2, and by admitting mercury adjust the volume of this again to 2 ml. Then the mercury would rise by $\frac{760 \times 0{\cdot}5}{2{\cdot}0}$ mm. = 190 mm. in the barometer. Suppose that now we admit by the tap G a small amount of a NaOH solution to absorb the CO_2, and again bring the volume to 2 ml. We should find the pressure had fallen to 171 mm. above the "wet level." Similarly, after absorbing the oxygen by hydrosulphite solution the pressure would fall to 133 mm. above the vacuous wet level.

---

[*] Cohen and Smith. *J. Biol. Chem.*, 1919, **39,** 489.

[†] Haldane. *J. Physiol.*, 1900, **26,** 497.

[‡] King *et al.* *J. Physiol.*, 1943, **102,** 21P ; *Lancet*, 1947, **253,** 789.

[§] Wang. *J. Biol. Chem.*, 1928, **77,** 409.

[||] Courtice and Douglas. *J. Physiol.*, 1947, **105,** 345.

[¶] Van Slyke and Stadie. *J. Biol. Chem.*, 1921, **49,** 1.

[**] Van Slyke. *J. Biol. Chem.*, 1927, **73,** 121.

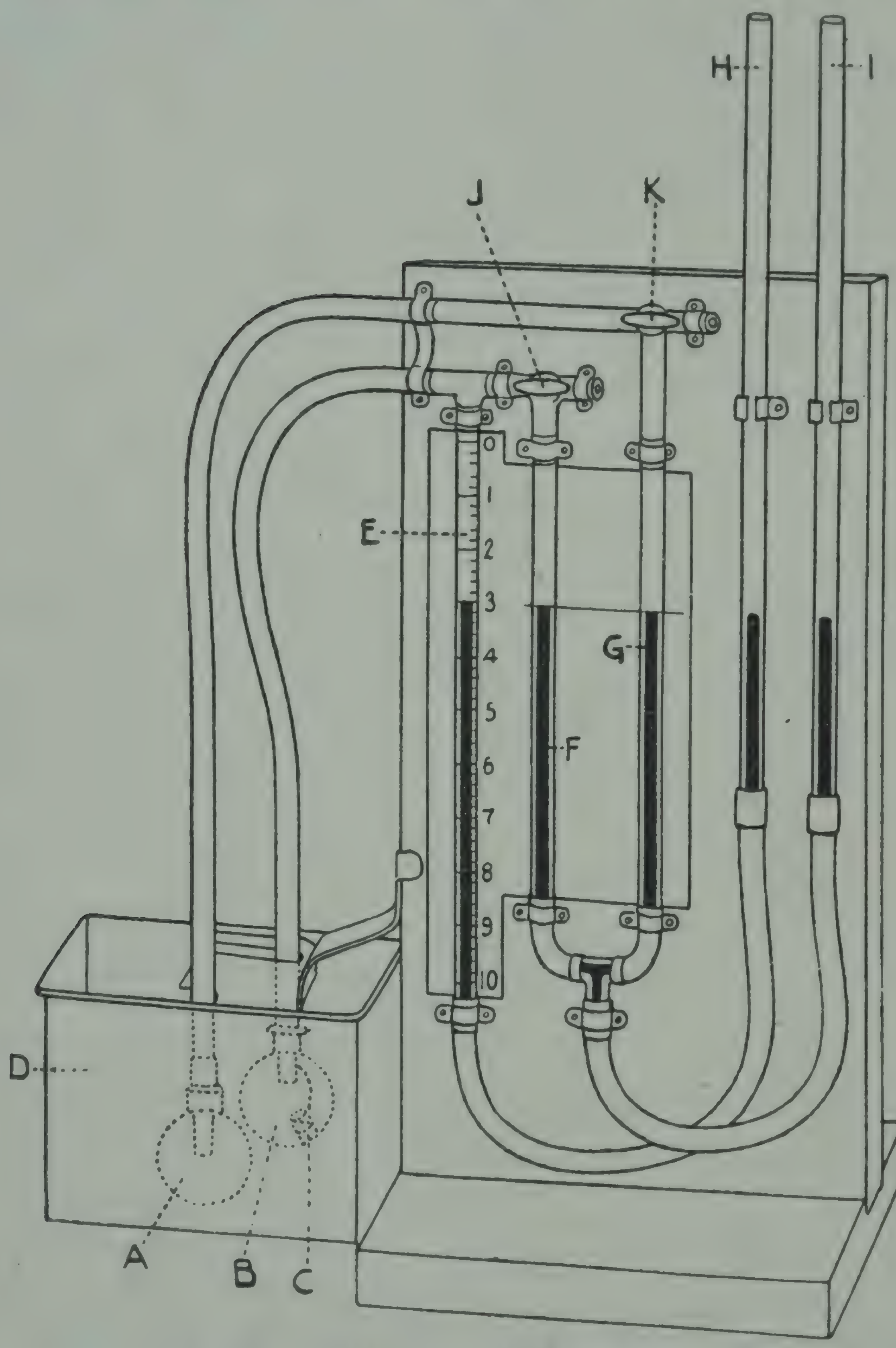


FIG. 325. Haldane's Blood-gas Apparatus.

Evidently the composition of the gas could have been deduced from these readings, and also the amounts at N.T.P. Thus the pressure of oxygen would be 171 − 133 = 38 mm., and its volume would be $\frac{2 \times 38}{760} = 0{\cdot}1$ ml., which we know to be correct.

In using the apparatus for the estimation of the gases in blood, an acidified ferricyanide solution is first introduced into the chamber, which is then evacuated and shaken by the motor attachment, to extract all dissolved air. A measured volume of blood (1 ml.) is then introduced through tap G, the chamber again evacuated and shaken, and thus all gas expelled from the blood, the CO_2 by the acid, the oxygen by the ferricyanide, and the nitrogen by the evacuation. The volume of gas is then brought to 2 ml., and the pressure p_1 is read. Then 1 ml. of N·NaOH is admitted to absorb CO_2, the volume again brought to 2 ml., and pressure p_2 is read. From p_1-p_2 with certain corrections, the volume of CO_2 is found. Next 1 c.c. of an alkaline hydrosulphite is introduced, volume brought to 2 ml. and p_3 read. Oxygen is calculated from p_2-p_3. For very small volumes of blood the gases may be read at 0·5 ml., so that pressures four times as great are given throughout. In practice various precautions have to be taken, but the calculations are greatly simplified by the use of appropriate tables.*

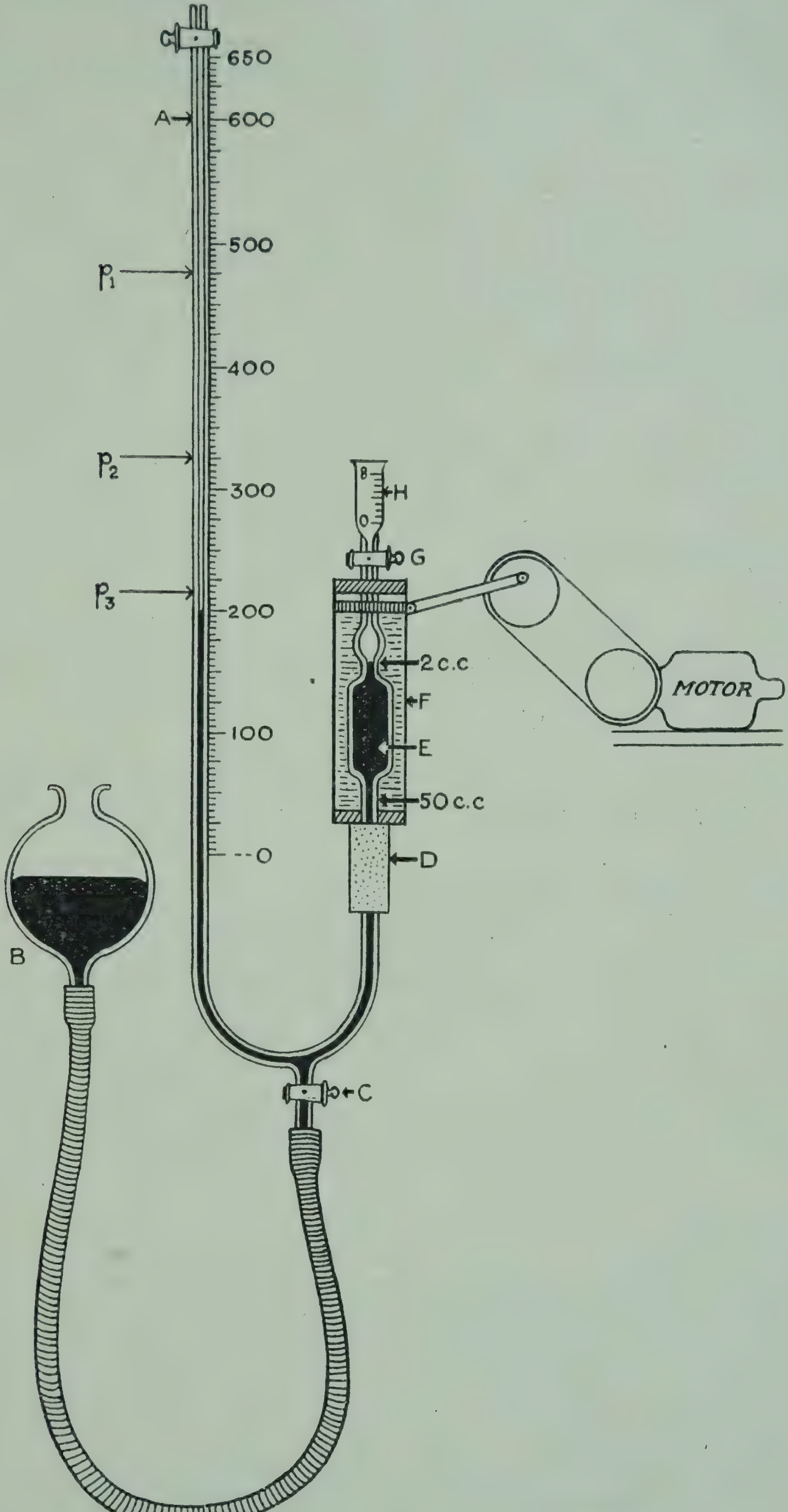


FIG. 326. Diagram to illustrate the principle of the Van Slyke manometric apparatus.

---

* PETERS and VAN SLYKE. Quantitative Clinical Chemistry, 1932. (Baillière, Tindall and Cox).

OXYGEN CONTENT AND PER CENT. SATURATION. It is often necessary to know the actual *content* of oxygen in the blood under given conditions, and also the degree of full saturation which that content represents. This is called the *percentage saturation* *i.e.*, $\frac{\text{content}}{\text{capacity}} \times 100$. Thus if a blood of 20 volumes per cent. oxygen capacity contains under given conditions 15 volumes per cent., then its degree of saturation is 75 per cent., its *deficit* of saturation being 25 per cent.

It is easy, on any given sample of blood, to determine at one time the oxygen content, percentage saturation and capacity, by the use of Haldane's apparatus. For this purpose, 1·5 ml. of borate buffer at *p*H 10·0 is placed in flask B; then 1 ml. of blood is run beneath the surface of the alkali, and the flask corked and immersed in the water bath D until its temperature is constant. The flask A contains 2·5 ml. of water, and is similarly corked, and the T-taps J and K left open to the air. The levelling tube I is adjusted to bring the water in F and G to the marks, and H to bring that in E to a suitable level, and then taps J and K are turned so that the T-bores lie thus : ⊣ The burette E is now read, and the flask B agitated for a few minutes so that the layer of blood is mixed with the alkali, laked, and fully saturated with oxygen. After again equalising temperatures, the levels in F and G are adjusted by the levelling tubes H and I. The blood has now absorbed the amount of oxygen necessary to saturate it ; this volume (" the deficit ") is read off on the burette E and corrected for temperature and pressure ; call it D c.c. Flask B is now opened, and 0·25 ml. of 6 per cent. ferricyanide, with a pinch of saponin, added to the tube C, the flask closed, and the oxygen capacity of the now fully saturated blood determined by adding the ferricyanide. Call it S ml. After making certain minor corrections * we can then calculate the original oxygen content C :

$$C = S - D$$

and the percentage saturation, P

$$P = \frac{S - D}{S} \times 100.$$

For instance, suppose that the oxygen capacity was 17·4 ml. per 100 ml. of blood and the deficit was 0·92 ml. per 100 ml. blood, then the blood as taken contained 17·4 — 0·92 = 16·48 ml. oxygen per 100 ml., while the percentage saturation would be $\frac{16\cdot48}{17\cdot40} \times 100 = 94\cdot7$.

The degree of oxygen saturation of the blood can also be approximately estimated by the use of a photoelectric method, and continuous records of the oxygen saturation of the circulating blood can thus be made.†

Average Red Cell Data for Human Blood. The data accepted in Britain for average human red cell count and hæmoglobin content at the beginning of the present century appear now to be too low.

Price-Jones ‡ found as follows for 100 healthy men and women in London :

	Men (average age 27)	Women (average age 24)
Red cells, p.c. mm.	5·428 million	5·012 million
Hæmoglobin, p.c. (Haldane) . . .	105·42	98·26
Colour Index $\left[= \frac{\text{Hb\%} \times 5}{\text{Millions of cells per c.mm} \times 100} \right]$	0·97	0·98

In the U.S.A., still higher Hb average values have been reported for young adult males. The following data, averaged from 115 healthy American male students, aged 18 to 30, are given by Foster and Raymond Johnson (1931) § (*v.* Fig. 324).

(1) Red cells = 5·26 million, p.c. mm.
(2) Hæmoglobin = 15·63 grams per cent.
(3) Oxygen capacity = 21·05 vols. per cent. (= 114 per cent. on Haldane Scale and Colour Index = 1·08).
(4) Hæmatocrit volume : 46·7 per cent. (Heparinised.)
(5) Average volume of corpuscle = $88\cdot8\mu^3$ per corpuscle $\left(= \frac{(4)}{100 \times (1)}\right)$.
(6) Average corpuscular Hb = $30\cdot03 \times 10^{-12}$ grams $\left(= \frac{(2)}{(1) \times 10^5}\right)$.
(7) Relative corpuscular Hb = 34·77 per cent. $\left(= \frac{(6)}{(5)}\right)$.

---

* HALDANE. " Respiration," p. 410. (Yale Press, 1922.)

† GOLDIE. *J. Scient. Insts.*, 1942, **19**, 23.

‡ PRICE-JONES. *J. Path. and Bact.*, 1931, **34**, 779.

§ FOSTER and RAYMOND JOHNSON. *Proc. Soc. Exp. Biol. Med.*, 1931, **28**, 929.

THE SPECIFIC GRAVITY OF THE BLOOD. The specific gravity of the blood may be determined by discharging drops of it into a liquid of known specific gravity, with which it will not mix. The simplest technique is to use a series of solutions of copper sulphate of different strengths and of densities covering the required range. The solution in which the drops remain suspended is that of the specific gravity of the blood.* From the specific gravity of blood and of plasma the Hb content, plasma protein, and hæmatocrit volume can be roughly inferred by reference to nomograms. The specific gravity varies between 1,055 and 1,066, being lowest in women. It is increased by loss of water, as after profuse perspiration, or by passive congestion of the part from which the sample is taken. It is also increased as a result of any operation upon a serous cavity in consequence of exudation of plasma in the inflamed or irritated part. It is diminished as the result of bleeding. The specific gravity of serum is 1,028 to 1,032, of corpuscles about 1,090. It is interesting to note that the specific gravity of the blood is highest in the fœtus at full term, when it amounts to 1,066, contrasting with that of the mother at the same time, the specific gravity of whose blood is only 1,050. The specific gravity rapidly falls after birth.

VISCOSITY OF THE BLOOD. Blood is more viscous than water, but since it is not a homogeneous liquid its apparent viscosity varies according to the method used for the measurement. Thus when viscosity is determined by the rate of flow along a capillary tube, the apparent viscosity is increased by increase in the relative number of corpuscles, by diminution in the bore of the capillary and by reduction in the driving pressure. A figure commonly given is five times that of water.

THE SUSPENSION STABILITY OF THE BLOOD. From the size of the red cells and the difference in specific gravity between them and the plasma, it can be calculated that when blood is allowed to stand the corpuscles should sink at a speed of about 0·2 mm. per hour ; yet the experience of centuries has shown that the speed with which sedimentation occurs is far more rapid and shows enormous variations, being greatly increased in some disorders.

In the days when blood-letting was almost universal in the treatment of disease, it was a frequent observation that on allowing the blood to stand there was formed at the surface of the clot a pale layer, known as the *buffy coat.* This was due to the fact that the corpuscles settled so rapidly that the clot was formed in part in the supernatant plasma, and was, therefore, pale in colour.

Fåhræus studied the sedimentation of citrated blood by allowing samples to stand undisturbed in long narrow tubes, and measuring the length of clear plasma after one hour. The following average results were obtained :—

New-born children	0·5 mm. per hour.
Normal men	3·3 ,, ,,
Non-pregnant women	7·4 ,, ,,
Pregnant women at term.	44·9 ,, ,,

The rapid rate of sedimentation is due to the phenomenon known as *agglutination,* which consists in the corpuscles aggregating together into clumps. In blood with a sedimentation value of 1 mm., each clump consists of about eleven corpuscles, while one with a value of 75 consists of clumps each composed of about 58,000 corpuscles. If the corpuscles and plasmata be separated off from two bloods with very different suspension stabilities, *e.g.* normal male and gravid female bloods, and then exchanged, it is found that the speed of settling is determined by some constituent of the plasma and not of the corpuscles. The logarithm of the maximal sedimentation velocity is proportional to the concentration in the plasma of the specific substances causing agglutination. This constituent is usually associated with the globulin fraction.

For Reference

FÅHRAEUS (1929). The Suspension Stability of the Blood. *Physiol. Rev.*, **9,** 241.
GARREY and BROWN (1935). Variations in White Blood Cell Counts. *Physiol. Rev.*, **15,** 597.
M. R. C. REPORT. Hæmoglobinometry and the Use of the Hæmatocrit. *B. M. J.*, Feb. 20th, 1943, p. 209.
PRICE-JONES *et al* (1935). Hæmatological Standards of Healthy Persons. *J. Path. and Bact.*, **40,** 503.

Blood Groups. The phenomenon of agglutination is often shown in an exaggerated form when the blood of one individual or species is added to the blood, or to the plasma or serum, of another. If, for instance, a drop of human blood be mixed, on a slide or white plate, with some horse serum, the agglutination is at once evident to the unaided eye. It is found that human beings fall into four main groups, when agglutination tests are performed.

Agglutination is due to the interaction between substances in the corpuscles, called *agglutinogens,* and substances called *agglutinins* in the plasma. The agglutinable substances, or agglutinogens, occur on the corpuscles in an insoluble form, and sometimes in tissue fluids in soluble form, and in various forms of mucin, and appear themselves to be mucoids.*

---

* PHILLIP, VAN SLYKE *et al. Bull. U.S. Med. Dept.*, Dec., 1943, **71,** p. 66.

The agglutinins are naturally occurring anti-bodies. Two main agglutinogens, called A and B, may exist in human red cells, and these can react with two agglutinins, called α and β, found in human plasma or serum. Thus α agglutinates A corpuscles and β agglutinates B. The cells of an individual may contain one or both agglutinogens, or none at all, and similarly for the agglutinins of the serum. Subjects thus fall into four groups, according as the cells are A, B, AB (*i.e.* with both) or O (*i.e.* with neither). A blood with A cells cannot have a plasma with α agglutinins, nor can one with B cells have β plasma, *i.e.* the agglutinogens have no antigenic properties. The blood with A cells contains the β plasma, that with B contains α, that with AB contains neither agglutinin, and that with the O group of cells contains both α and β agglutinins in the plasma. Hence when the sera of one group are mixed with cells of other groups the results are as shown in the table (+ indicates agglutination).†

Corpuscle Group	Serum contains	Effect of adding Serum of Blood Group			
		A (serum = β)	B (Serum = α)	AB (Serum = o)	O (Serum = αβ)
A	β	−	+	−	+
B	α	+	−	−	+
AB	o	+	+	−	+
O	αβ	−	−	−	−

The corpuscles of Group O are not agglutinated by the plasma of any group, and Group O are accordingly called "Universal Donors," since their blood is safe for transfusion into any person. In Group AB the plasma is *o*, and does not agglutinate any corpuscles, so that people belonging to this group could receive blood from any healthy person. The other groups are as shown. Thus Group B could act as a donor to Groups AB or B only. For transfusion work, account need only be taken of the action of the agglutinins of the recipient on the cells of the donor ; the serum of the donor will be so diluted that its effect on the recipient's corpuscles is relatively slight.

It is usual to carry out grouping tests by the use of tested sera belonging to Groups A and B. Drops of the diluted citrated blood of the person to be tested are mixed on a white plate with a drop of each of these, and, as in the table :

Blood of Group AB is agglutinated by serum of A and B (containing β and α respectively).
,, Group A ,, ,, B (containing α).
,, Group B ,, ,, A (containing β).
,, Group O ,, ,, neither.

Among English people, about 2 per cent. belong to Group AB, 40 to Group A, 15 to Group B, and 43 to Group O. With other races the distribution is different.‡ The agglutinogens are passed on by inheritance, A and B being Mendelian dominants, the agglutinins and therefore Group O, being recessives —a fact of medico-legal importance in cases of disputed paternity,§ *e.g.* if the mother is, say, B, and her child A, the father must have been A or AB.

The classification of blood groups is made still more precise, especially in examination of medico-legal cases, by the discovery that, quite independently of the AB system mentioned above, another classification can be made, based on different agglutinogens called

---

* MORGAN. *Experientia.*, 1947, **3**, 257.

† The above is the League of Nations grouping. Other groupings often used are that due to Moss, whose Groups I, II, III and IV correspond respectively to AB, A, B and O of the above and that of Jansky, largely used in U.S.A., where, in the same order, the groups are numbered IV, II, III, and I.

‡ MILLOT. *Antiquity*, 1935, **9**, 399.

§ HARLEY and ROCHE LYNCH. *B.M.J.*, 1937 (i), 163 ; THOMAS. *Ibid.* (ii), 315.

M and N. These are identified by special immune rabbit sera, and provide three independent types, M, N and MN. As these occur independently of the AB groups, the combination yields twelve possible blood types, and in fact, yet other subsidiary groups are known.

An important factor, called the Rh factor was discovered in 1940. It is an antigen of the Rhesus monkey, is present in the red cells of 85 per cent. of white people, and may be transmitted from father to child. To take the simplest case, if the mother is Rh negative, and the fœtus Rh positive because the father is positive, she forms anti-bodies to it ; these pass by the placenta to the fœtus and cause extensive blood destruction (erythroblastosis fœtalis) in the new-born child, which is shown by extra medullary blood formation and persistence of erythroblasts in the blood. This state of affairs is probably the commonest cause of fœtal death. A further difficulty is that the woman who has become sensitive will always remain so, and will herself be thenceforth unable to support a transfusion with a Rh positive blood. The inheritance of the Rh factor is too complex to discuss here, since it can be subdivided into several (probably at least six) genetic sub-groups.

Blood Transfusion. The facts of blood grouping acquire great practical importance, because agglutination is the precursor of hæmolysis, and because the plasma of one human being frequently causes agglutination of the corpuscles of another. Transfusions of incompatible blood may lead to severe symptoms and may be fatal. The incompatible corpuscles are hæmolysed in the body, hæmoglobin is excreted in the urine, or the kidneys may even cease to function, and a state of shock and collapse ensue. In carrying out the urgent operation of blood transfusion, it is therefore a matter of great importance to ensure that the corpuscles of the donor will not be agglutinated by the plasma of the recipient, *i.e.*, that the recipient's plasma does not contain agglutinins that react with the donor's corpuscles. To make quite sure of this, it is necessary, if circumstances allow, to test out the blood of the donor directly against the serum of the recipient, in addition to the grouping test, because some corpuscles may in spite of identity of main groups, belong to subsidiary groups and prove to be incompatible with a given serum. The M and N factors do not complicate the test because there are no natural agglutinins in man corresponding to them, but occasional anti-Rh antibodies may occur. Further essential precautions include the avoidance of clotting, and of infection, and the introduction of blood at the proper temperature, and at the proper rate, into the veins of the recipient. Coagulation is usually prevented by the addition of sodium citrate, 3·2 per cent. solution, to five times its volume of blood. The citrated blood can be preserved at 5° C. for periods up to ten days, or longer, if an equal volume of 5·4 per cent. glucose should be added to the mixture. Collections of stored blood (blood banks), duly grouped are often kept at large hospitals. When hæmolysis begins to set in rapidly the stored blood must be rejected.

The blood donor usually gives 420 c.c. of blood at each bleeding : this results in a loss of about 2·3 g. of Hb per 100 c.c. blood, and it takes about 50 days for complete restoration, though if iron is taken by mouth the time for return to normal is reduced considerably.* Men regenerate Hb at about 0·05 g./100 c.c./day, and women at about 0·04 g./100 c.c./day.

For Reference

Cappell (1946). The Blood Group Rh. *B.M.J.*, Oct. 26, 1946, p. 601 ; Nov. 2 1946, p. 641.

Doan (1927). The Transfusion Problem. *Physiol. Rev.*, **7**, 1.

M.R.C. (1946). No. 9 The Determination of Blood Groups. H.M. Stationery Office.

M.R.C. (1948). No. 19. The Rh Blood Groups. H.M. Stationery Office.

Strandskov (1944). Five Human Blood Characteristics. *Physiol. Rev.*, **24**, 445.

Wiener (1943). Blood Groups. (Springfield.)

The Reaction of the Blood. The blood is slightly alkaline to litmus, and can be titrated with an acid. Such a procedure, however, gives a measure, not of the degree of alkalinity, but of the alkaline reserve—*i.e.* of the total amount of base in combination with weak acids which can be replaced by a stronger acid. This *alkaline reserve* consists almost exclusively of sodium bicarbonate.

In Van Slyke's method the alkaline reserve of the plasma is determined by finding out how much CO_2 is evolved from 1 c.c. of the oxalated plasma, when this is treated with 0·5 c.c. of 5 per cent. sulphuric acid. Since the amount of carbon dioxide taken up depends on the partial pressure of the CO_2 in the atmosphere to which the plasma is exposed, the plasma is first shaken up with alveolar air provided by the experimenter himself, which always contains about 5·5 per cent. of CO_2.

Normal human blood plasma treated in this way yields between 60 and 70 c.c. of CO_2 per 100 c.c., which corresponds to about 0·2 per cent. of sodium bicarbonate.

The *hydrogen ion concentration* of the blood can be determined by the electrometric method, or by the use of indicators, as described on p. 65.

---

* Fowler and Barer. *J. Amer. Med. Assoc.*, 1942, **118**, 421.

By these methods the hydrogen ion concentration of the blood, *i.e.* of the plasma, at 38° is found to be $0{\cdot}4 \times 10^{-7}$, or pH 7·4. It might be thought that with such a feeble alkalinity the merest trace of acid added to the blood would suffice to make it acid. It is found, however, that a relatively large proportion of an acid must be added to the blood in order to produce a considerable change in its reaction. This is due to the fact that the sodium bicarbonate acts as a "buffer"—*i.e.* a substance which can take up acid or alkali with a minimal change of reaction. As in other buffer systems (*v.* p. 64) the hydrogen ion concentration depends on the ratio of the weak acid to its salt, *i.e.* to $\frac{H_2CO_3}{NaHCO_3}$. This in blood is usually $\frac{1}{20}$ when the molar concentrations are taken, and the *p*H of the blood can be determined by finding the concentrations of free and fixed CO_2, and the use of the *Henderson-Hasselbalch equation :*

$$pH = pK_1 + \log \frac{[NaHCO_3]}{[H_2CO_3]}$$

where pK_1 has the value of 6·1 (*i.e.* $pH = 6{\cdot}1 + \log 20 = 7{\cdot}4$ approx.) The property of the blood of retaining a nearly constant reaction, even though fixed acids are added to it, is of immense importance, since all cellular functions are acutely sensitive to changes in reaction, and as we shall see later, the activity of the respiratory centre is largely dependent on the hydrogen ion concentration of the blood with which it is bathed.

The alkaline reserve of the blood is significant, since pronounced diminution indicates in all probability the production of fixed acids in the tissues, and a progressive reduction will precede the point at which the "buffer" action of the sodium bicarbonate is lost, and the blood then responds to any addition of acid by a considerable change in reaction. It is only when the alkaline reserve has been greatly reduced that a true condition of *acidæmia*, or increased H-ion concentration, can come into being.

For Reference

HENDERSON (L. J.) (1928). Blood. (Yale and Oxford.)

LOVATT EVANS (1930). Recent Advances in Physiology, 4th ed. (Chap. V. The Reaction of the Blood). (Churchill.)

THE OSMOTIC PRESSURE OF THE BLOOD. Since the blood serves as a circulating medium, by means of which the composition of the tissue juices is maintained constant, its osmotic pressure must be of considerable importance. The osmotic pressure of the blood depends on its molecular concentration and can be determined by any of the methods mentioned earlier (p. 58), *e.g.* by the determination of the freezing-point. The depression of freezing-point, Δ, of mammalian blood in equilibrium with air is about 0·53°. This is equal to that of a 0·9 p.c. (= 155 milli-equivalents per litre) sodium chloride solution, which is therefore often taken as isotonic with the blood.

By the use of Hill's thermopile method for vapour pressure, Magaria * found the average osmotic pressure of the blood of adult men to be equal to that of 0·945 per cent. NaCl ($\Delta = 0{\cdot}553°$ C.) when the blood was in equilibrium with 5 per cent. CO_2. The blood of women had a lower osmotic pressure = 0·927 per cent. NaCl ($\Delta = 0{\cdot}543°$ C.). Under normal conditions the osmotic pressure is remarkably constant ; it is reduced slightly by large ingestions of water and increased slightly during severe muscular exercise.

Actually the osmotic pressure of the contents of the red corpuscles is a little lower than that of the plasma, but the difference is not great enough for hæmolysis appreciably to alter the freezing-point or osmotic pressure of the blood.

For frog's blood Hill and Kupalov found an osmotic pressure = 0·725 per cent. NaCl.

THE ELECTRICAL CONDUCTIVITY OF THE BLOOD. In a solution it is only the dissociated ions which have the power of carrying electric charges. The conductivity of blood serum is therefore determined almost entirely by its content in salts. Since this is approximately constant, the conductivity of serum varies within very narrow limits. The conductivity of blood varies, however, within wide limits. The corpuscles present a resistance to the passage of the charged ions and therefore of the electric current through them, so that the larger the number of corpuscles contained in a given specimen of blood the lower will be the conductivity of the latter. Stewart made use of this fact as a basis for a method of determining the relative volume of corpuscles and plasma. Owing to the fact that the red cells carry a negative charge, they move towards the anode when an electric current is passed through the blood.

---

* MAGARIA. *J. Physiol.*, 1930, **70**, 417

THE GENERAL COMPOSITION OF THE BLOOD

Human blood contains from rather over one-third to one-half of its weight of corpuscles. It contains from 20 to 25 per cent. of solids.

The following table shows a typical percentage composition of human blood :—

Normal Human Blood

	Per cent. in Whole Blood	Per cent. in Plasma	Per cent. in Corpuscles
Total solids	22·0	9·3	34·6
Inorganic solids	0·75	0·8	0·7
Serum albumin	2·8	4·6	—
Serum globulin	0·95	2·2	—
Fibrinogen	0·25	0·4	—
Hæmoglobin	15·6	0·0	33·0
Non-protein nitrogen	·032	·025	·044
Urea nitrogen	0·012	·012	·011
(= urea)	0·026	—	—
Amino nitrogen	0·006	0·005	0·008
Creatinine	0·002	—	—
Creatine	0·005	—	—
Uric acid	0·003	—	—
Reducing power (as glucose)	0·10	0·09	0·10
True glucose	0·07	0·08	0·06
Lactic acid	0·020	0·025	0·014
Pyruvic acid	0·0008	—	—
Total fatty acids	0·36	0·38	0·35
Total "lipoids"	—	0·67	—
Fat	0·10	0·14	0·07
"Lecithin"	0·3	0·2	0·4
Cholesterol	0·21	0·22	0·19
Total acetone bodies (as acetone)	0·001	—	—
Sodium	0·20	0·33	0·065
Potassium	0·20	0·015	0·38
Calcium	0·006	0·01	0·001
Magnesium	0·003	0·0025	0·0035
Iron	0·052	0·0001	0·10
Chlorine (chloride)	0·29	0·36	0·20
Phosphorus, total	0·045	0·010	0·079
Phosphorus, total acid soluble	0·030	0·0035	0·060
,, inorganic	0·005	0·003	0·006
,, "lipoid"	0·013	0·007	0·018
Sulphate sulphur	0·0005	0·001	—
Bicarbonate (as HCO_3)	—	0·15	—
,, (as $NaHCO_3$)	—	0·21	—

Blood Plasma is resolved by clotting into serum and fibrin. The fibrin forms only 0·2 to 0·4 per cent. of the total weight of blood. The *serum* contains, in 100 parts, 8 to 9 parts of solids, of which 7 to 8 parts consist of proteins, while the salts make up to about 1 part.

The chief salt present in the serum is sodium chloride, and next to this comes sodium bicarbonate, and besides these two we find traces of potassium, sodium and calcium, and phosphates. Ammonia is absent from fresh blood, but after shedding appears in amount up to 0·1 mg./100 c.c. This is formed by deamination of adenylic acid derivatives.*

The anions and cations of the plasma are often expressed in terms of milli-equivalents per litre. The total base of the serum in man is kept remarkably constant at the isotonic value of 155 milliequivalents per litre. The distribution of ions is roughly as follows :—

---

* Conway and Cooke. *Bioch. J.*, 1939, **33**, 457.

	Cations. m.eq./l.		Anions. m.eq./l.
Na =	142.	Cl' =	103
K =	5.	HCO_3' =	28
Ca =	5.	HPO_4'' =	2
mg =	3.	Protein′ =	16
		SO″ and organic =	6
	155		155

The distribution of the anions is more variable than that of the cations. but the sum of anions is normally constant and equal to that of the cations.

Traces of fats, cholesterol (about ⅓ free and ⅔ as esters), lecithin, glucose, urea, amino-acids and other nitrogenous extractives are constantly found in the serum. Tiny particles called *hæmokonia* or chylomikrons are always visible in plasma under dark-ground illumination. They consist of fat. The fats are much increased after a meal rich in them and may give the serum a milky appearance. The yellow colour of the plasma is due to the presence of small amounts of bilirubin, carotene (C_{40} H_{56}) and xanthophyll ($C_{40}H_{56}O_1$).

The Proteins of the Plasma. Dried plasma, and various protein fractions of plasma have found extensive use in the clinic, as blood substitutes, etc. The plasma contains a number of different coagulable proteins. We may divide them into those which are the precursors of fibrin and those which remain in the serum.

FIBRINOGEN is a globulin converted on coagulation into fibrin. It has a high molecular weight (500,000) and its molecules are large and elongated (900 × 33 Å). It is formed in the liver. This is proved by the fact that after excision of the liver the fibrinogen content of the blood slowly declines.* If the circulating blood is defibrinated (by removing it, whipping and re-injecting) the fibrinogen content is restored in a few hours ; no restoration occurs if the liver has been previously removed, but rather a continued loss.† The plasma contains normally about 0·25 per cent. (0·19 to 0·38 per cent.) of fibrinogen.

Prothrombin is also a globulin (p. 529), which is concerned in fibrin formation.

Fibrin. Fibrin is easily obtained by whipping blood as it flows from the vessels with a bundle of wires or twigs, and then washing the stringy threads so obtained (see p. 527). Fibrin is insoluble in water and in dilute salt solutions. It slowly dissolves in 5 per cent. sodium chloride, but it is probable that its solution is effected by the agency of minute traces of proteolytic enzymes adherent to the fibrin as it is precipitated. In dilute acid, such as 0·2 per cent. hydrochloric acid, fibrin swells into a clear jelly which very slowly undergoes solution with the formation of acid metaprotein and proteoses.

THE PROTEINS OF THE SERUM. The serum proteins are generally grouped in two classes : the serum albumins and the serum globulins. All the proteins are precipitated by saturation with ammonium sulphate, but only the globulin by half-saturation. The normal range for albumin in human plasma is 3·7 to 4·7 per cent. and for globulin (plus fibrinogen)1·9 to 3·6 per cent. The ratio of globulin to albumin is known as the "protein quotient". It varies in different animals, but in the same individual it is almost constant in the blood, serum, lymph and transudates, though the total amounts of protein in these may be very different. Globulins are probably the more important because they include (in the γ-fraction) the various anti-bodies, such as precipitins, antitoxins, lysins, agglutinins, etc., which play a large part in the phenomena of immunity (*v.* Chapter XXXVIII). The serum proteins are formed in the liver ; the albumin provides 80 per cent. of the colloid osmotic-pressure of the plasma ; they also act as a reserve of protein for the body in time of need.

Serum albumin is a mixture of proteins of relatively small molecular weight (69,000) and size (150 × 38 Å). It can be precipitated by complete saturation with ammonium sulphate. The fraction also contains a glyco-protein called globoglycoid.*

---

* NOLF. *Arch. internat. de physiol.*, 1905, **3**, 1.

† MCMASTER and DRURY. *Proc. Soc. Exp. Biol. Med.*, 1928, **26**, 490 ; JONES and SMITH *Amer. J. Physiol.*, 1930, **94**, 144.

The *serum globulins* are mixtures of proteins of greater molecular weight and size (*e.g.* γ-globulin mol. wt. 150,000 and size 320 × 36Å). They can be divided into three fractions, namely, *euglobulin I*, insoluble in water at *p*H 7, *euglobulin II*, insoluble in water at *p*H 6 (precipitated by CO_2), and *pseudo-globulin*, "soluble" in water. Pseudo-globulin is, however, not soluble in *complete* absence of salts, and is therefore thrown down by electro-dialysis.

Hardy † showed that with acids or alkalies the globulin forms "salts" which ionise in solution so that in an electric field the protein ions move with a speed dependent on their weight, size and shape. According to these speeds of migration, they have been grouped into α, β and γ fractions. The usual proportions are $\alpha = 0{\cdot}84$, $\beta = 0{\cdot}78$, $\gamma = 0{\cdot}66$ g./100 ml.

THE CONDITION OF THE PROTEINS IN THE BLOOD SERUM. Although it is easy to separate one or more different forms of protein from serum, we have evidence that these proteins in the serum are present in a somewhat different form. When examined by the ultra-centrifuge there are found to be less of the large molecules resembling globulin than would be expected from salting-out, and correspondingly more of the smaller molecules of albumin dimensions.‡ On dilution this distribution is altered until finally it agrees with the proportions found on analytical separation.

Serum can be readily filtered through a porous cell, and this method is adopted for obtaining it free from contamination by micro-organisms; yet globulin when isolated in solutions will not pass through a porous cell.

In explanation of these and other facts, it is now considered that the serum proteins consist of numerous closely related substances, separable roughly into the various groups, *e.g.* albumin, the globulins, etc., but each member of which is in fact an association of mutually interacting systems of colloids.§ By simple means, such as dialysis, dilution or addition of salt, fractionation by alcohol, electrophoresis at given *p*H, etc., these complexes can be broken up into various fractions.

For Reference

COHN (1941). The Properties and Functions of the Plasma Proteins. *Chem. Rev.*, **28**, 395.

EDSALL (1947). The Plasma Proteins and Their Fractionation. *Adv. in Protein Chemistry*, Vol. III, p. 384. (Acad. Press Inc. N.Y.)

FOLIN (1922). Non-protein Nitrogen of the Blood in Health and Disease. *Physiol. Rev.*, **2**, 460.

GRAM (1924). Composition and Physical Properties of Normal Human Blood (a compilation of values from the Literature). *Amer. J. Med. Sci.*, **168**, 511.

GUEST and RAPOPORT (1941). Organic Acid-Soluble Phosphorus Compounds of the Blood. *Physiol. Rev.*, **21**, 410.

MADDEN and WHIPPLE (1940). Plasma Proteins. *Physiol. Rev.*, **20**, 194.

MYERS (1924); and MYERS and MUNTWYLER (1940). Chemical Changes in the Blood and their Clinical Significance. *Physiol. Rev.*, **4**, 274; *Physiol. Rev.*, **20**, 1.

WEECH (1938). The Significance of the Albumin Fraction of Serum. (*Harvey Lects.*, 1938–9; p. 57.)

THE RED BLOOD CORPUSCLES

The red blood corpuscles, or *erythrocytes*, in man are non-nucleated bi-concave discs, which, when in the body, are about 8·8 μ in diameter and about one-third of this in thickness. In the dried, fixed film they are, owing to shrinkage, on the average only about 7·2 μ in diameter, extreme values being between about 6·7 and 7·7 μ. The actual size-distribution in normal people is shown in Fig. 327, which illustrates the narrow range of variation in size.‖ The *average* size of the red cells can be determined in a film, fresh or dried, by a diffraction method,¶ using an instrument called an eriometer or halometer.

The red corpuscles are not of rigidly fixed size when in the body. The circumstance which is chiefly responsible for change of size in the body is alteration in the reaction of the blood plasma; reduction of the alkalinity of the

---

* HEWITT. *Bioch. J.*, 1938, **32**, 26.

† HARDY. *J. Physiol.*, 1905, **33**, 251.

‡ MCFARLANE. *Bioch. J.*, 1935, **29**, 407, 660, 1175, 1209.

§ RIMINGTON. *Ergebn. d. Physiol.*, 1933, **35**, 712.

‖ PRICE JONES. *J. Path. and Bact.*, 1929, **32**, 479.

¶ PIJPER. *S. Afr. Med. Rec.*, September 27th, 1919; ALLEN and PONDER, *J. Physiol.*, 1928, **66**, 37.

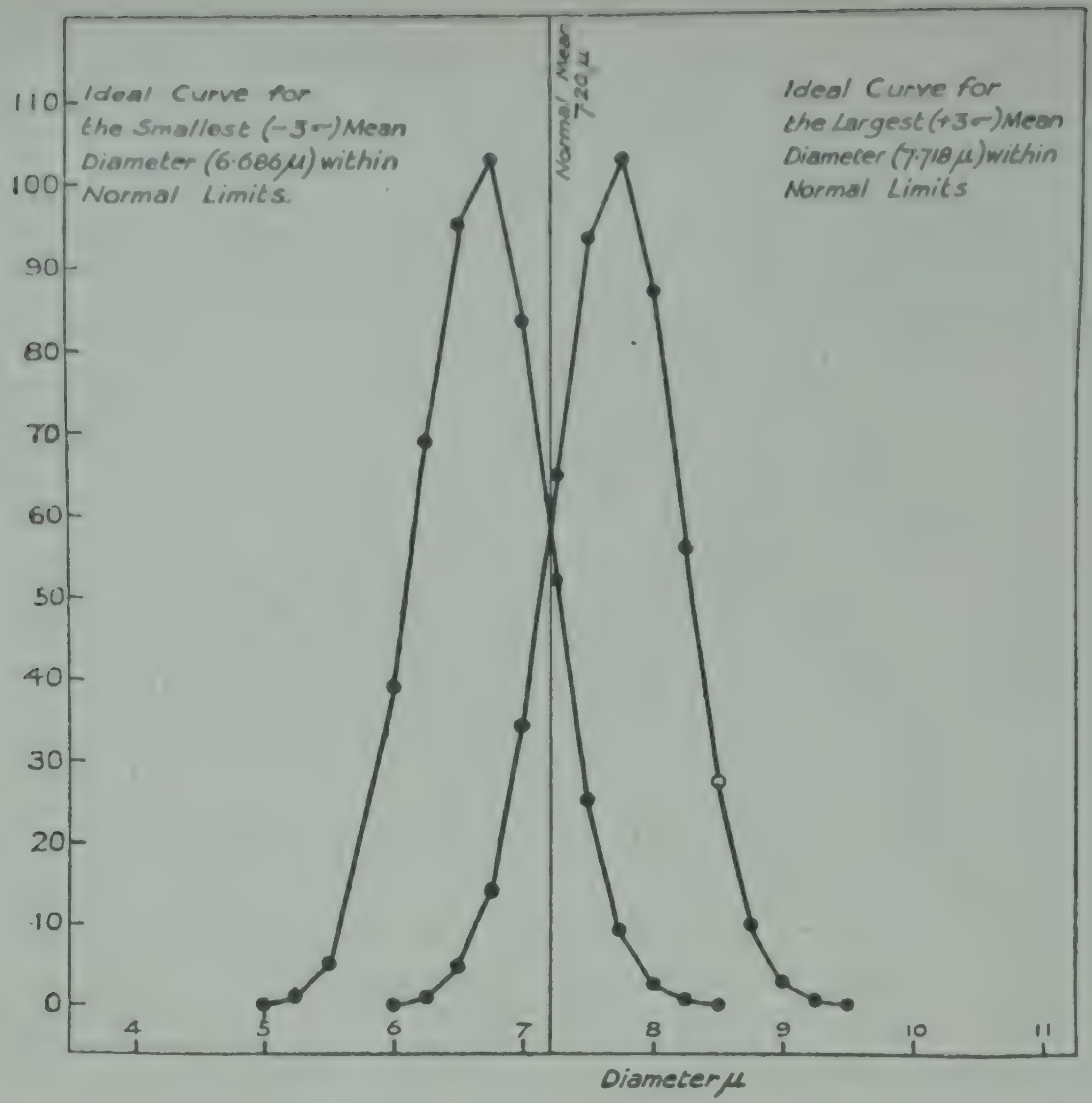


Fig. 327. Size-distribution curves of normal erythrocytes showing the extreme range for smallest and largest mean diameters (500 cells). The curve for average normal individuals is midway between these two. (Price Jones.)

plasma causes swelling of the corpuscles, and *vice versâ*. Hence it follows that the corpuscular volume is slightly greater in venous than in arterial blood.

The colour of a single corpuscle when viewed under the microscope is yellow, the red colour being apparent only when larger numbers are seen together. The red corpuscles form about 50 per cent. of the total mass of the blood. They are soft, flexible and elastic, so that they can readily squeeze through apertures and canals narrower than themselves, without undergoing permanent distortion. When examined by the usual histological methods, or by dark-ground illumination, the normal erythrocyte appears quite structureless. Each red corpuscle has been described as consisting of a framework or stroma, composed chiefly of protein and lipide material, containing in its meshes a red colouring matter, hæmoglobin, to which is due the colour of the corpuscles and of the blood itself. No limiting envelope is histologically demonstrable, though a convenient view is that a monomolecular layer of some material—probably lipides, such as lecithin—exists at the surface.

In nearly all mammalia the red corpuscles are of the character described (Fig. 328). In the camel they are oval in shape, but otherwise resemble the corpuscles of other mammals. In all other classes of vertebrata the red corpuscles are oval, nucleated cells; the hæmoglobin is diffused through the protoplasm of the cell body and does not extend to the nucleus. During the early part of fœtal life the corpuscles of mammals are also nucleated, but in the adult condition the erythrocytes, except under abnormal conditions, lose the nucleus before entering the blood

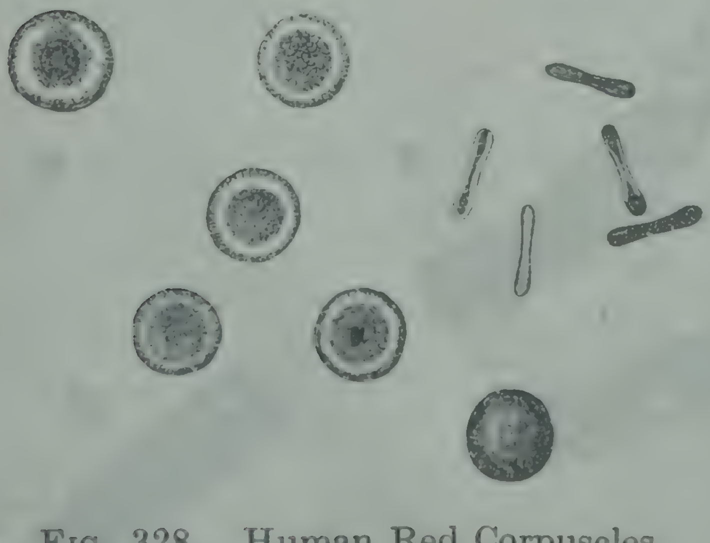


Fig. 328. Human Red Corpuscles. On the right of the figure the corpuscles are seen on edge. (Swale Vincent.)

stream. The small size and great number of the red corpuscles determine that a very large area of surface of red corpuscles is exposed to the plasma. This great extent of surface is of importance in facilitating the exchange of material, especially oxygen, between the corpuscle and the surrounding plasma.

The volume of each corpuscle has been estimated as about 90 μ^3 (= ·000000090 c. mm.), and its surface as ·000128 mm.^{2}, so that the total surface of red corpuscles in the blood of a man weighing about 70 kilos (assuming his total blood as $\frac{1}{13}$ of the body weight) would be about 3,000 sq. metres, or 1,500 times the surface of the body itself.

Bürker has shown that in various species there is a close relation between the surface area of the red cell and its hæmoglobin content, the ratio being constant at about 32 × 10^{-14} g. of hæmoglobin per square μ of corpuscle surface.

Osmotic Relationships of the Red Corpuscle. If the blood plasma be concentrated by evaporation or by the addition of neutral salts, its osmotic pressure rises, and water diffuses from the corpuscles into the plasma to equalise the osmotic pressure within and without the corpuscle. The latter therefore becomes wrinkled or *crenated*. On the other hand, dilution of the plasma diminishes its osmotic pressure to below that of the corpuscles, and water therefore passes into the latter, which swell up and become spherical ; and, if the plasma be made sufficiently dilute, the corpuscles liberate the hæmoglobin they contain, leaving the colourless swollen *stromata* floating in the plasma. At the same time the blood becomes transparent and darker in colour, and is called " laked " or hæmolysed blood. This is called *hæmolysis*. In partial hæmolysis the corpuscles do not all burst ; microscopical examination of the process, with dark-ground illumination, shows that many corpuscles merely swell to a great size, the " ghosts ", which remain after hæmolysis but are nearly invisible to ordinary observation. They become more visible, and return to their former size, if the solution is made isotonic again. We may perhaps attribute the ultimate escape of the contents to the fact that the cell contains insufficient lipide to form a continuous layer over so large a surface.

Blood may be laked by any of the following means :—

(*a*) Addition of a small amount of ether or other lipide solvent.
(*b*) Free dilution with water, or dialysis.
(*c*) Alternate freezing and thawing of the blood.
(*d*) Addition of bile salts or of saponin.
(*e*) The action of foreign blood serum or of various hæmolysins whose nature we shall discuss later.

The erythrocytes of different individuals, especially in certain pathological states, exhibit different amounts of resistance to hæmolysis. As a measure of the " fragility " or ease with which hæmolysis may be induced by hypotonic solutions, we may plot a curve showing per cent. hæmolysis at different concentrations of NaCl * and find the concentration at which 50 per cent. hæmolysis occurs. In normal blood this is at 0·334 to 0·398 per cent. NaCl : in hæmolytic jaundice it may be as high as 0·7 to 0·8 per cent. ; after splenectomy fragility is reduced to 0·25 to 0·35 per cent. Cells of venous blood are slightly more fragile than those of arterial blood, and in general an increase in the ratio thickness/diameter of the cells increases fragility. Fragility is increased by change of pH above or below 8·1 and by cold.† The corpuscles of normal mammalian blood in presence of air neither gain nor lose volume in a solution containing 0·9 per cent. sodium chloride. The osmotic pressure as determined by the freezing-point of such a solution is identical with that of the blood, which in most mammals has a freezing-point of about — 0·53° C. When the CO_2 is prevented from escaping from human blood, its osmotic pressure is nearer to that of 0·95 per cent. NaCl (men), or 0·93 per cent. NaCl (women). The salt solution which is normal for frogs' blood contains only 0·725 per cent. sodium chloride.

---

* Creed. *J. Path. and Bact.*, 1938, **46**, 331.
† Dacie and Vaughan. *Ibid.*, 1938, **46**, 341.

The behaviour of the red corpuscle, when immersed in various solutions, suggests that its external layer is impermeable to most cations and neutral salts as well as to cane sugar and glucose, but this is not strictly true. We can, however, make isotonic solutions with sodium chloride, sodium sulphate, potassium nitrate, or cane sugar. On the other hand, a solution of urea behaves towards the corpuscles like distilled water. If some red corpuscles be added to a 1 per cent. solution of urea in normal salt solution, they neither shrink nor swell, but if added to plain aqueous urea solutions of any strength the corpuscles will be laked. There are many substances to which the corpuscles are permeable, *e.g.* alcohol, chloroform, ether, etc. In their permeability the corpuscles resemble most other vegetable and animal cells in permitting the passage of those substances that are soluble in lipides, which are invariable constituents of all living cells.

For Reference

Ponder (1924). The Erythrocyte and the action of the Simple Hæmolysins. (Oliver and Boyd, Edinburgh and London.)

Ponder (1936). The Kinetics of Hæmolysis. *Physiol. Rev.*, **16**, 19.

Ponder (1948). Hemolysis and related Phenomena. (London : Churchill.)

Chemistry of the Red Blood Corpuscles. The chief constituent (95 per cent.) of the solids of the red cells is the chromoprotein, hæmoglobin, which is present chiefly as a potassium salt. It is to this substance that the red colour of blood is due. The other 5 per cent. corresponds to the stroma, consisting of stroma protein (nucleo-protein), lecithin, cholesterol and salts. There is a striking contrast between the salts of the corpuscles and those in the plasma, the former consisting, in man, chiefly of potassium salts, *e.g.* hæmoglobinate, diphosphoglycerate, the latter of sodium chloride, which in some animals is entirely wanting in the corpuscles.

This is due to active excretion of $\mathrm{Na}^{\cdot}$ by the corpuscle, $\mathrm{K}^{\cdot}$ then being absorbed into it passively to maintain the cation concentration. When shed (citrated) blood is stored, as for blood transfusion, at 2° to 4° C. the K of the plasma rapidly rises to 8 to 10 times its initial value, because of loss of K from the cells when the " sodium pump " ceases to act : incubation with glucose restores the normal relation by providing energy to restart it.

The total base concentration in the red cells is greater than that in the serum, and the anions are correspondingly less, in accordance with the Donnan equilibrium (p. 83), which tells us that :—

$$\frac{\text{anions (cells)}}{\text{anions (plasma)}} = \frac{\text{cations (plasma)}}{\text{cations (cells)}} = \frac{\mathrm{H}^{\bullet}\text{ (plasma)}}{\mathrm{H}^{\bullet}\text{ (cells)}} = \frac{\mathrm{Cl}'\text{ (cells)}}{\mathrm{Cl}'\text{ (plasma)}}, \text{ etc.}$$

The Donnan ratio $\frac{\text{anions (c)}}{\text{anions (p)}}$ is usually about 0·75.

The non-diffusible anions to which this phenomenon is due, are hæmoglobin, diphosphoglycerate, and adenosine triphosphate : the last two are the chief sources of acid-soluble phosphorus in the corpuscles.

The blood corpuscles contain a greater percentage of solids than any soft tissue of the body. The blood corpuscles have 37 per cent. solids, as against muscular tissue with 25 per cent. or nervous with 22 per cent. solids. Hence the cells contain at least 30 per cent. hæmoglobin, which is far beyond its solubility. Some form of combination therefore exists in the corpuscles, which keeps the hæmoglobin from separating out in crystalline form.

In investigations on the chemistry of the red cells it is usual first to remove the plasma by centrifugalisation and washing. The blood, which has been defibrinated, or prevented from clotting by the addition of oxalate, is centrifuged until all the formed elements are thrown down. The sediment is then mixed with normal saline and again centrifuged, and this process repeated twice in order to wash away adherent plasma or serum. After the final washing the paste of corpuscles can be laked ; for this, two volumes of distilled water saturated with ether are added to one volume of caked corpuscles. The fluid is once more centrifuged in order to throw down white blood corpuscles. A 1 per cent. solution of acid sodium sulphate is now added drop by drop until the solution acquires the opaque appearance presented by ordinary blood. On centrifuging, the stromata are thrown down and can be collected and washed with distilled water several times on the centrifuge. Removal of stromata from laked cell solutions can also be effected by shaking for an hour or two with asbestos pulp and then removing this, with adherent stromata, by centrifugalisation. The cells also contain about 0·3 per cent. lipides, which is all in the stroma. About 10 to 15 per cent. of the dry weight of stroma material consists of lipides, of which 60 per cent.

is phospholipide (half of it cephalin), 30 per cent. free cholesterol and 10 per cent. cholesteryl esters and fats.*

From a laked solution of corpuscles, oxyhæmoglobin can be obtained in a crystalline form with varying readiness according to the animal from which the blood is derived. Thus in the case of the rat, the guinea-pig, the dog and the horse, it is sufficient merely to cool the laked blood, to about − 10° C., in order to obtain a large crop of hæmoglobin crystals. Crystallisation is facilitated by the addition of 25 per cent. of alcohol to the mixture, but this partly denatures the hæmoglobin. To avoid this, the following simple method depends on the fact that oxyhæmoglobin is less soluble than reduced hæmoglobin. Washed horse corpuscles are dialysed in collodion sacs against the osmotic pressure of the hæmoglobin. When fully laked, *débris* is removed by centrifugalisation, and then the cooled solution well oxygenated, when oxyhæmoglobin crystals separate out. These can be collected, suspended in distilled water, and reduced *in vacuo*, when they dissolve ; on re-oxygenation the solution again crystallises.†

Properties of Hæmoglobin. The crystals thus obtained are as a rule microscopic in size. Most animals yield an oxyhæmoglobin which crystallises in rhombic prisms or needles belonging to the rhombic system, though there are exceptions to this (Fig. 329). The solubility of the crystals differs according to the animal from which they have been derived. They are more soluble in alkalies than in water and contain carbon, hydrogen, nitrogen, oxygen, sulphur and iron. The chief differences between various animals appears to have relation to the sulphur, and to the varying avidity with which the hæmoglobins combine with oxygen.‡ All specimens are alike in containing a constant proportion of iron, viz. 0·335 per cent. It is in the *ferrous* state.

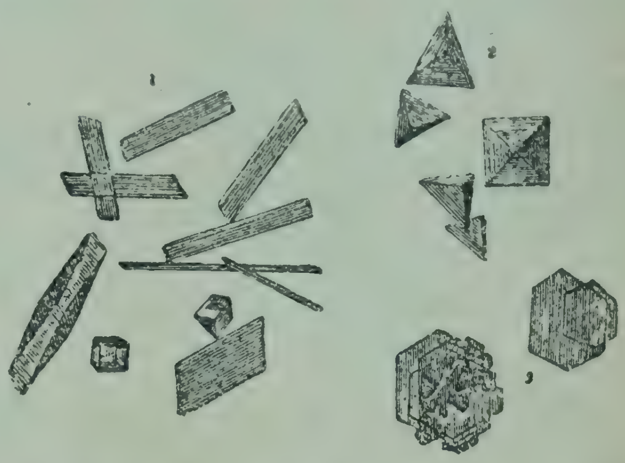


Fig. 329. Crystals of Oxyhæmoglobin. 1. From rat. 2. From guinea-pig. 3. From squirrel.

Hæmoglobin readily combines with free oxygen to form *oxyhæmoglobin*,§ and it is found that for each atom of iron two atoms of oxygen can be taken up, *i.e.* 393 c.c. of oxygen per gramme of iron present in hæmoglobin. This corresponds to an oxygen capacity of about 1·3 c.c. oxygen per gramme of hæmoglobin, if we take 17,000 as the molecular weight of Hb containing one atom of iron. The figure usually taken is 1·34 c.c. O_2 per g. Hb, but it is so difficult to prepare hæmoglobin pure that this figure is provisional. Oxyhæmoglobin is thus a compound, in definite proportions, of oxygen and hæmoglobin. It can be easily dissociated and is split up by such simple means as exposure to a vacuum (*v.* Chapter XXXIX) ; the fluid then gives off oxygen and the colour changes from a scarlet to a dull bluish red. The same change can be effected by treating a solution of oxyhæmoglobin with reducing agents such as hydrazine, sodium dithionite (hydrosulphite) ($Na_2S_2O_4$), or ammonium sulphide. The oxygen in oxyhæmoglobin can be replaced by equivalent quantities of some other gases. Thus if carbon monoxide be led through a solution of oxyhæmoglobin, oxygen is given off, and its place is taken by an equal volume of carbon monoxide with the formation of a more stable compound, carbon monoxide hæmoglobin, or *carboxyhæmoglobin*. This body is dissociated only with extreme slowness and is unaffected by the addition of reducing agents. By using special precautions to prevent oxidation of the gas, the carbon monoxide can be replaced in this compound by nitric oxide, NO. We have therefore a series of three compounds which can be arranged in order of diminishing stability, thus :

NO-hæmoglobin.
CO-hæmoglobin.
O_2-hæmoglobin.

---

* Erikson, *et al.* *J. Biol. Chem.*, 1938, **122**, 515.

† Dudley and Lovatt Evans. *Bioch. J.*, 1921, **15**, 487.

‡ McCarthy. *J. Physiol.*, 1935, **86**, 77.

§ Normal blood may contain up to 5 per cent. of its Hb in an inactive form, incapable of carrying oxygen : the nature of this inactive Hb is unknown.

The poisonous properties of carbon monoxide are chiefly due to its power of turning out the oxygen from the oxyhæmoglobin, thus depriving the tissues of the oxygen which is normally carried to them by the red corpuscles. Small amounts of CO-Hb are found in the blood of smokers.

Another compound of interest is *sulphæmoglobin* which is formed by the addition of H_2S to reduced hæmoglobin. Sulphæmoglobin is formed in the body when excessive intestinal putrefaction leading to formation of H_2S occurs, especially if certain drugs such as sulphonamide are being administered. The resulting condition of sulphæmoglobinæmia may be dangerous.

Hæmoglobin and its derivatives give well-marked absorption spectra. Thus dilute solutions of oxyhæmoglobin show two well-marked absorption bands between Fraunhofer's lines D and E. The centre of the band nearest to D corresponds to λ 579, and is often spoken of as the α band, while the second band, the one next to E, called the β band, is broader, has less sharply defined edges, and its centre corresponds approximately to λ 542. On concentrating the solution or using thicker layers, a point is reached at which the two bands fuse into one, and with a still stronger solution the whole of the spectrum is absorbed with the exception of the red end.

If a reducing agent be added to the solution of oxyhæmoglobin, the two bands disappear and their place is taken by a more diffuse band lying midway between the two (Fig. 330), its centre corresponding to λ 559. This is the absorption spectrum of hæmoglobin (*reduced hæmoglobin*).

The spectrum of carboxyhæmoglobin is very similar to that of oxyhæmoglobin, the α band being slightly ($\lambda = 570\ m\mu$) nearer to the violet end. In a mixture of oxy- and carboxyhæmoglobin, the shift of the α band is proportional to the fraction of the hæmoglobin present as the carboxy-compound, a fact which has been utilised by Hartridge for the determination of carboxyhæmoglobin by means of his reversion-spectroscope.* It is of a brighter red than oxyhæmoglobin. The fact that CO-hæmoglobin is not altered by reducing agents can be shown by adding sodium dithionite to CO-hæmoglobin, when no change is observed. The amount of CO present in a blood may be estimated chemically by expelling all the blood gases by acidified ferricyanide solution, and subsequent analysis of the gases, by the Van Slyke apparatus.†

All these derivatives of hæmoglobin, besides their absorption bands in the visible spectrum, have characteristic absorption of light in the ultra-violet spectrum. In the case of oxyhæmoglobin this absorption causes a band (Soret's band) which occupies the greater part of the spectral region between Fraunhofer's lines G and H with its centre at $\lambda = 415\ m\mu$. In reduced hæmoglobin this band is displaced towards the visible part of the spectrum.

Another derivative of hæmoglobin is *methæmoglobin.* This is normally only present in traces in the blood. It may be prepared by the addition of a ferricyanide, permanganate, or certain other oxidising agents to solutions of oxyhæmoglobin. It is a chocolate-brown substance, crystallisable, and gives an absorption band in the red between Fraunhofer's lines C and D ($\lambda = 634\ m\mu$). It is unaltered by exposure to a vacuum. It contains no available oxygen, but may contain a hydroxyl group. Its iron is in the *ferric* state.‡ Dithionite reduces this to the ferrous state, giving reduced hæmoglobin, from which, by shaking with air, oxyhæmoglobin can again be obtained. When methæmoglobin is formed from oxyhæmoglobin the whole of the loosely bound oxygen of the latter is given off, a fact which is utilised for the estimation of the oxygen held in combination in the blood.

The formation of methæmoglobin is shown in the approximate equation :

$$HbO_2 + K_3Fe(CN)_6 + KOH \longrightarrow HbOH + K_4Fe(CN)_6 + O_2$$

---

* Hartridge. *J. Physiol.*, 1912, **44**, 1.

† Sendroy and Liu. *J. Biol. Chem.*, 1930, **89**, 133 ; Peters and Van Slyke. "Quantitative Clinical Chemistry," 1932, 328. (Baillière, Tindall and Cox.)

‡ Conant and Scott. *J. Biol. Chem.*, 1926, **69**, 575.

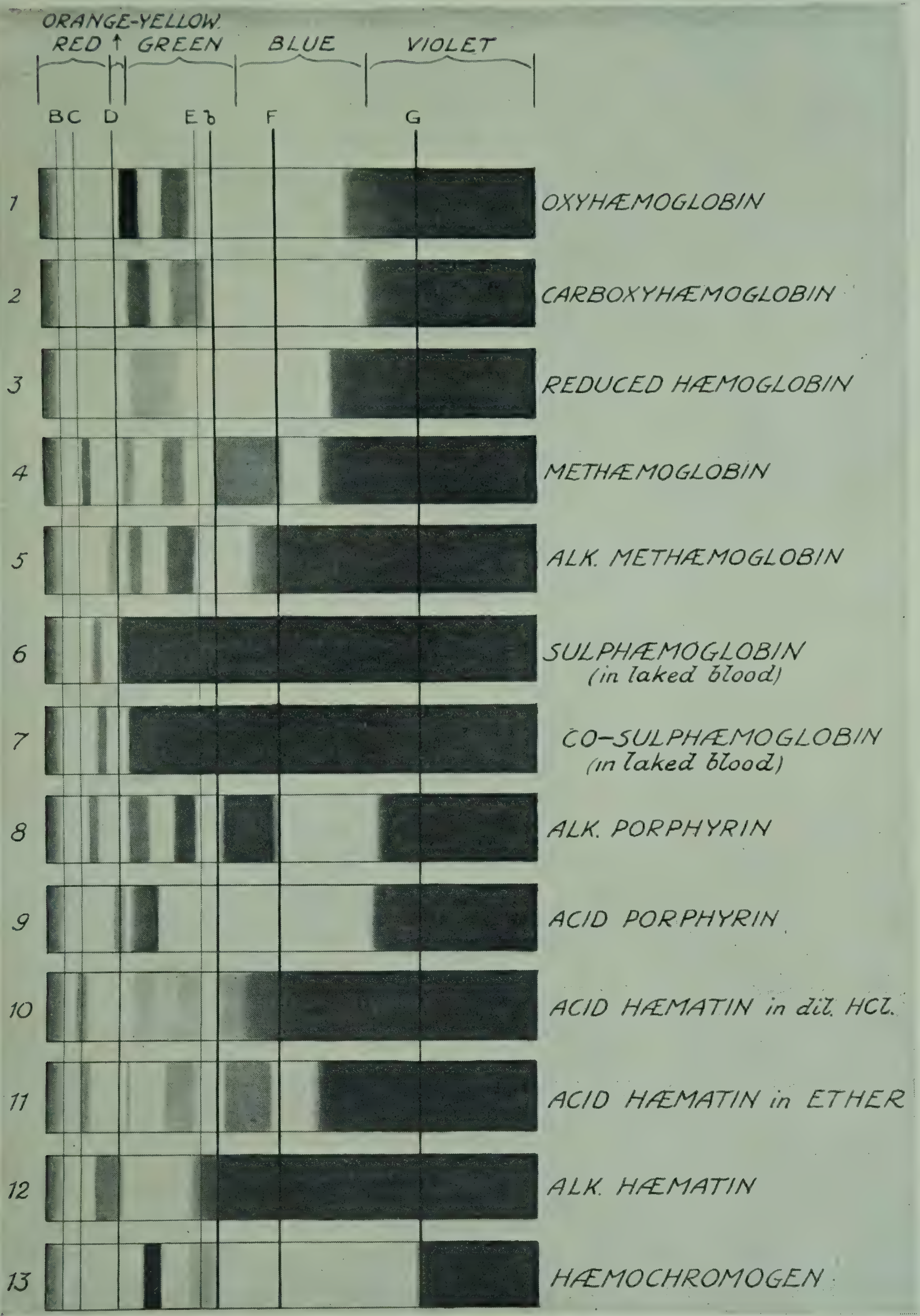


FIG. 330. Absorption spectra of hæmoglobin and its derivatives.

(From HARRISON's "Chemical Methods in Clinical Medicine." Churchill.)

[To face p. 512.

in which it is seen that the hæmoglobin molecule is converted into methæmoglobin by change of the iron to the ferric state, the ferricyanide being simultaneously reduced to ferrocyanide.

By adding a cyanide to methæmoglobin the latter is converted into cyanhæmoglobin (spectrum resembling reduced Hb); by spectroscopic observation of absorption of light at 630 mμ before and after adding cyanide, the amount of methæmoglobin in a blood may be determined.*

DERIVATIVES OF HÆMOGLOBIN. Hæmoglobin is a compound of an iron-containing coloured group, called *hæm*, with a protein which varies somewhat in different animals. The prosthetic group is identical in every case where it has been examined. A separation of the prosthetic group from the protein moiety can be effected with extreme ease, and occurs whenever the hæmoglobin is treated with weak acids, with alkalies, or is heated above 70° C. The protein group is known as *globin*.

It is a strongly basic protein, classified with the histones. It is by the formation of co-ordination linkages between the histidine groups of the globin and the N of the hæm ring that the union is effected. Unless prepared with great care the globin from blood is obtained in a denatured condition. Undenatured globin was prepared by R. Hill and Holden by breaking up oxyhæmoglobin solution with the smallest possible amount of dilute HCl at − 2° C., adding cold water, ether and kieselguhr and shaking well. It is then filtered and kept at − 2° overnight. Metaprotein is thrown down by adding ammonia, and is filtered off. The filtrate is freed from ether by vacuum treatment and the salts removed by dialysis.

Hæmoglobin yields about 94 per cent. of globin and about 4·5 per cent. of the chromogenic group, *hæmatin*. In order to obtain hæmatin in a pure condition, it is usual to start with the crystalline derivative *hæmin*. When some dried blood is heated with a crystal of common salt and glacial acetic acid on a slide, the residue contains reddish brown hæmin crystals (Fig. 331), and by a large-scale modification, hæmin may be prepared in quantity. It is a hydrochloride of hæmatin of the formula: $C_{34}H_{32}O_4N_4ClFe$ (Fischer). The iron is in the ferric state. By reduction in alkaline solution it yields the ferrous compound, *hæm* ($C_{34}H_{32}O_4N_4Fe$). By dissolving hæmin in alkalies and adding excess of acid, hæmatin is precipitated. It is insoluble in water, alcohol or ether. but is slightly soluble in glacial acetic acid and in absolute alcohol. It is easily soluble in concentrated sulphuric acid, with decomposition, losing its iron and being transformed into *hæmatoporphyrin*. The formula of hæmatin is probably $C_{34}H_{32}O_4N_4Fe.OH$. Its compounds with acids and alkalies are spoken of as acid and alkaline hæmatin, and each gives a characteristic absorption spectrum (Fig. 330).

Hæmochromogens. In the presence of certain foreign matters, alkaline hæmatin, when treated with reducing agents, forms a substance known as hæmochromogen. The same change is observed when crude alkaline hæmatin, made by the action of alkalies on ordinary blood, is reduced by ammonium sulphide. These facts receive an explanation from the discovery by Anson and Mirsky † that the hæmochromogens are co-ordination compounds of reduced hæmatin with one of many possible nitrogenous compounds. When made from blood, the nitrogenous compound is a denatured globin, but other proteins, or nicotine, hydrazine, pyridine or ammonia can take its place, to yield very similar hæmochromogens. When hæmochromogen is prepared from blood, the changes would be:

$$\text{Oxyhæmoglobin} + \text{alkali} \longrightarrow \text{Hæmatin} + \text{Globin (denatured)}$$
$$\text{Globin (den.)} + \text{Hæmatin} \,[+ \text{reducing agent}] \longrightarrow \text{Hæmochromogen.}$$

When prepared with other compounds, *e.g.* pyridine, the hæmochromogen is analogous, *e.g.* a compound of reduced hæmatin and pyridine. When reduced hæmoglobin is treated with alkali, the globin is denatured, but not

---

* EVELYN and MALLOY. *J. Biol. Chem.*, 1938, **126**, 655.

† ANSON and MIRSKY. *J. Physiol.*, 1925, **60**, 50.

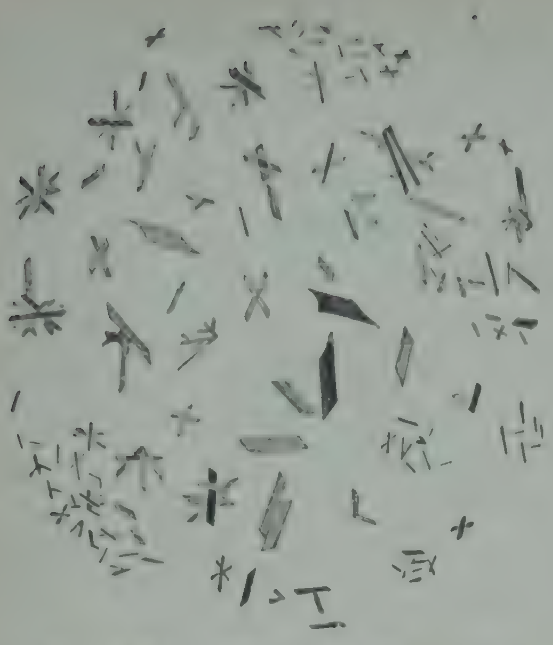


FIG. 331. Hæmin Crystals.

split off, and hæmochromogen results at once. It is evident then that "hæmochromogen" differs from hæmoglobin essentially in the fact that its protein is denatured, and in existing only in the reduced form; when oxidised it yields alkaline hæmatin and the protein or other nitrogenous body is split off.

The "blood" hæmochromogens in solution have a cherry-red colour, and when sufficiently diluted show two well-marked absorption bands (Fig. 330) which are situated between D and E; that nearest to D has very sharply defined borders; the positions of the two absorption bands are λ 555 to 559 and λ 525 to 528. The band nearest D is given by 1 in 25,000 hæmochromogen solutions, so that it is an even more delicate test for blood than is the spectrum of oxyhæmoglobin itself. Hæmochromogen does not combine with oxygen, but does so with carbon monoxide. When CO-hæmoglobin is treated with alkali in the absence of oxygen, the CO-hæmochromogen formed contains exactly the same volume of CO in combination as the original CO-hæmoglobin.*

Cytochrome. Hæmochromogens are found widely distributed in almost all tissues of animals and plants, and one series has been given the name cytochromes by Keilin. They exist in the oxidised and reduced forms, the latter of which shows a characteristic spectrum (Fig. 332). In living tissues examined by the microspectroscope the formation of the reduced cytochrome can be witnessed when oxygen is excluded, and it has been shown that cytochrome-*c* plays an important part in the execution of oxidative changes in the cells. When oxygen is excluded, the tissues reduce it; this is a result of the action of dehydrogenases, the cytochrome-*c* acting as hydrogen acceptor; when oxygen is readmitted, in presence of an enzyme, cytochrome oxidase, it is oxidised again and so it acts as an oxygen carrier. The amount of cytochrome in a tissue is closely related to the intensity of its respiration; moreover, when cyanide is added to the tissue, the cytochrome remains in the reduced state even when oxygen is allowed access to it, and at the same time the oxygen utilisation of the tissue is stopped. This is because cyanide paralyses the cytochrome oxidase.

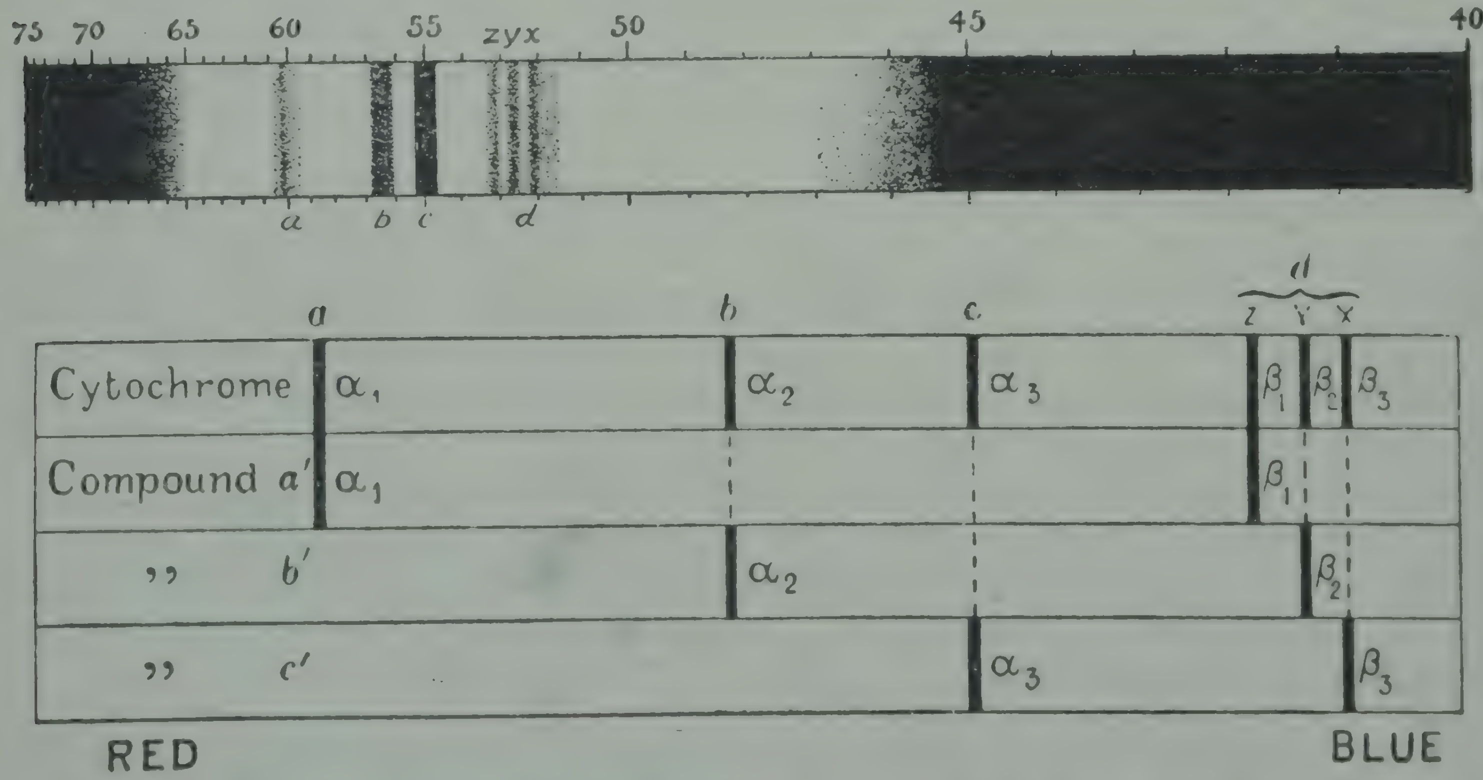


FIG. 332. *Above:* Absorption Spectrum of Cytochrome in Muscle of Bee. Reduced form as obtained by crushing tissue between two glass slips.
Below: Showing three of the separate spectra which together make up the above spectrum of mixed cytochrome. (Expanded scale.) These components and their bands, have now been re-named. Cytochrome *a′* now includes α and α_3, while *b′* and *c′* are called *b* and *c*. (After KEILIN.)

---

* ANSON and MIRSKY. *J. Physiol.*, 1925, **60**, 65.

Tissue-cytochrome is a mixture of four cytochromes, called α, α_3, b, and c of different properties* (Fig. 332).

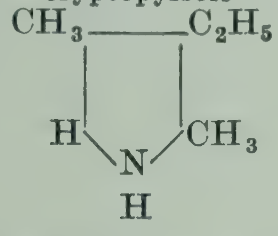
Hœmatoporphyrin. If hæmoglobin, hæmatin or hæmin be treated with a strong acid, it dissolves, forming a purple-red solution. On pouring this solution into a large quantity of water, hæmatoporphyrin is thrown down as a brown precipitate.

Hæmatoporphyrin is easily soluble in alkalies and somewhat less readily so in acids forming alkaline and acid hæmatoporphyrin respectively. Solutions of alkaline hæmatoporphyrin show four absorption bands.

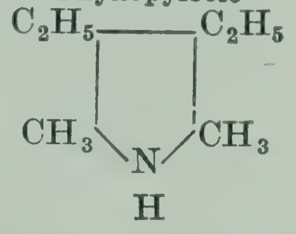
In the formation of hæmatoporphyrin from hæmatin, the iron of the latter has been split off by the action of the strong acid. The splitting off of iron occurs much more readily in the absence of oxygen. Porphyrins related to hæmatoporphyrin occur in minute quantities in normal urine and in larger quantities in certain toxic conditions, especially in poisoning by sulphonal, when the urine may have a red colour.

Chemical Relationships of Hœm Pigments. If hæmin or hæmatoporphyrin be reduced by the action of hydriodic acid we get a mixture of substituted pyrroles formerly known as hæmopyrroles. The hæmopyrroles, according to Willstätter, are three in number and have the following formula :

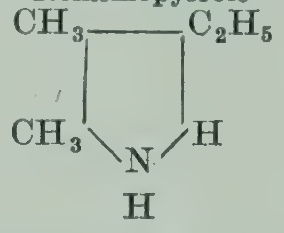
Cryptopyrrole



Phyllopyrrole

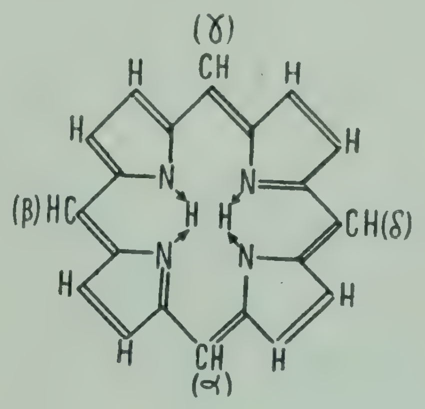


Isohæmopyrrole



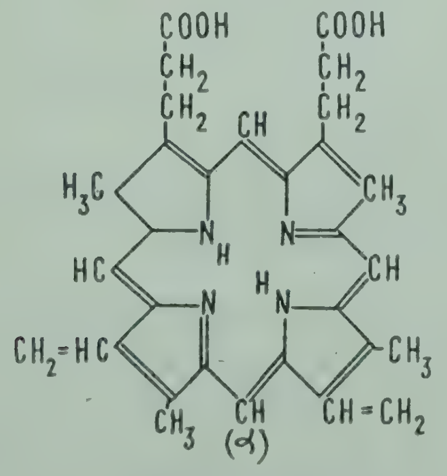
The same substances can be obtained from the chlorophyll of plants.

PORPHYRINS. Thus the same group forms the basis both of the substance which is responsible in the plant for the assimilation of carbon from carbon dioxide, and of the pigment which in the animal is the carrier of oxygen between the tissues and the surrounding medium. These and various other pigments found in animals and plants are derived from the chemical group of *porphyrins*, which consist of four substituted pyrrol rings. They are all derived from the parent substance *porphin* :—



The porphyrins of importance in the body are those known as I and III. These differ in their substituent groups. Hæmatoporphyrin may be taken as a type of III. Many porphyrins are known and some have been synthesised.

A possible structural formula for hæmatoporphyrin is :—



Hæmatoporpyrin.

---

* KEILIN and HARTREE. *Proc. R. S.*, 1939, 127B, 167.

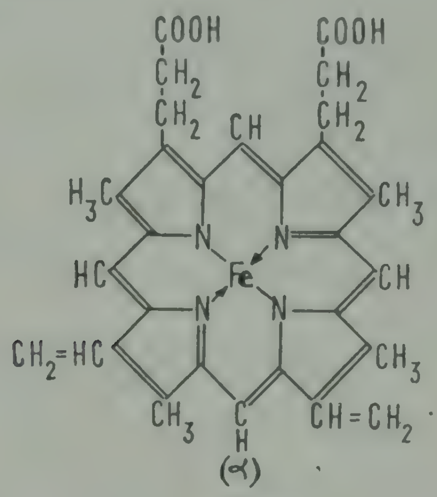
Derivatives of hæmatoporphyrin called uroporphyrins and coproporphyrins are found in urine and fæces respectively. They may belong to series I or III.

The relations of the chief simple porphyrins may be roughly shown as follows, designating the porphoryl group as Por :—

Protoporphyrin	Por : H_2
Uroporphyrin	Por : $(COOH)_6 . H_2$
Coproporphyrin	Por : $(COOH)_2H_6$

HÆMES. The porphyrins exist in various metallic derivatives called *metallo-porphyrins ;* those with iron form the group called *hæmes*, of which hæmatin and its derivatives are examples ; all the porphyrins form the corresponding ferri-porphyrins when treated with ferric chloride and sodium acetate. Thus protoporphyrin gives hæmatin.

The following is a suggested formula for hæm, the ferro-porphyrin group of hæmoglobin



Hæm.

Of the iron porphyrins there are two series, according to whether the iron is ferrous, or ferric :—

Ferro-porphyrins—	
Reduced hæmatin (hæm)	Por : Fe$^{\cdot\cdot}$
Reduced hæmoglobin	(Globin) (Por : Fe$^{\cdot\cdot}$)
Oxyhæmoglobin	(Globin) (Por : Fe$^{\cdot\cdot}$)O_2
Hæmochromogen	(Denatured globin) (Por : Fe$^{\cdot\cdot}$)
Ferri-porphyrins—	
Hæmatin	Por : Fe$^{\cdot\cdot\cdot}$OH
Hæmin	Por : Fe$^{\cdot\cdot\cdot}$Cl
Methæmoglobin	(Globin) (Por : Fe$^{\cdot\cdot\cdot}$OH)
Cyanhæmoglobin	(Globin) (Por : Fe$^{\cdot\cdot\cdot}$CN)

The power of loosely combining with O_2, CO and NO is confined to the ferrous compounds. Cyanhæmoglobin is, on the other hand, a CN derivative of a ferric compound, *e.g.* of methæmoglobin and is not formed by addition of HCN to oxyhæmoglobin.

Metallo-porphyrins containing metals other than iron are also known to occur naturally, *e.g.* turacin, a feather pigment containing copper ; and compounds containing cobalt, nickel, zinc, etc., have been artificially prepared. Again, iron porphyrins in which the porphyrin part is different, *e.g.* the respiratory pigment chlorocruorin, found in certain worms, have also been subjected to study.

The Synthesis of the Blood Pigments. Many of the porphyrins can be made synthetically, and H. Fischer and his pupils succeeded in synthesising hæmatoporphyrin, hæmatin and hæmin. Given hæmatoporphyrin, moreover, Laidlaw * showed that the iron could easily be introduced by adding ferrous sulphate and ammonia : in the presence of hydrazine hydrate hæmochromogen was formed, and from this hæmatin.

Laidlaw also succeeded in synthesising turacin, by treating hæmatoporphyrin with ammoniacal copper solution, showing that it is a compound corresponding to hæmatin, in which the place of iron is taken by copper.

The combination of globin and hæmatin to form hæmoglobin was effected by R. Hill and Holden.† They prepared globin from oxyhæmoglobin without denaturation, as described above ; the globin after purification was mixed with hæmatin in sodium carbonate, when methæmoglobin resulted. This on reduction gave reduced hæmoglobin, which on shaking with air gave oxyhæmoglobin, recognisable by its absorption spectrum. If denatured globin is employed, hæmochromogen is formed instead of methæmoglobin.

---

* LAIDLAW. *J. Physiol.*, 1904, **31**, 464.

† HILL and HOLDEN. *Bioch. J.*, 1926, **20**, 1326.

For Reference

Anson and Mirsky (1930). Hemoglobin, the Heme pigments and Cellular Respiration. *Physiol. Rev.*, **10**, 506.

Barcroft (1924). The Significance of Hemoglobin. *Physiol. Rev.*, **4**, 329.

Barcroft (1925). The Significance of Hemoglobin in Submammalian Forms of Life *Physiol. Rev.*, **5**, 596.

Barcroft (1928). The Respiratory Function of the Blood, Pt. II. Hæmoglobin. (Cambridge.)

Dobriner and Rhoads (1940). The Porphyrins in Health and Disease. *Physiol. Rev.*, **20**, 416.

THE LIFE-HISTORY OF THE RED BLOOD CORPUSCLES

Before Birth. At a very early stage of development there arises from the mesoderm in various situations the primitive connective tissue, or *mesenchyme*, and it is in this tissue, or from its direct descendants, that both blood cells and vessels are developed.

The first appearance of blood cells is in mesenchyme tissue situated outside the embryo itself, in the *area opaca* of the yolk sac, and in the amnion and chorion. Here are formed the blood islands, groups of cells in which the peripheral mesenchyme cells form the endothelium of the vessels, an intercellular fluid, the plasma, and the more centrally placed cells, the *primitive blood cells*, all of which latter are nucleated cells which closely resemble the parent mesenchyme cells. In man the earliest vessels appear in this way on the undersurface of the yolk sac and gradually extend over its surface, finally establishing connection with the similar vessels of the chorion. By the third week communication has been established between this system and another system of vessels, namely the blood vessels and heart of the embryo itself. The vessels of the embryo are likewise developed from the mesenchyme, but with the difference that in these vessels the differentiation of the mesenchyme is into *endothelium* and plasma only, with practically no primitive blood cells. The first blood cells of the embryonic circulation, therefore, represent an invasion of primitive blood cells from the yolk sac vessels into the vascular system of the embryo, when connection has been established between the two systems.

The primitive blood cells undergo further differentiation, and from them are derived cells called *hæmocytoblasts*. These cells are the primitive stem cells from which, according to some authorities, all the cells of the blood are ultimately derived. In the formation of the red cells, these cells are converted into the immediate precursors, called *erythroblasts*, and these, by the formation of hæmoglobin, into the fully formed red cells (Fig. 333).

At later stages of fœtal life, erythrocytes are produced, not only in the bone marrow, but also in the liver, spleen and lymph nodes.

Origin after Birth. In the fœtus the cells reach the circulation in an immature form and are nucleated, but soon after birth all the corpuscles are of the non-nucleated type. They are now formed, normally, only in the red marrow of the bones, and can be shown to be derived from nucleated erythroblasts by a process of degeneration and solution of the nucleus. We have definite proof that a continual formation of red corpuscles is proceeding throughout the whole of adult life. In an adult the total number of corpuscles remains approximately constant. By bleeding an animal we can diminish the total amount of corpuscles. The first effect of such a bleeding is that the fluid parts of the blood are made up, so that the volume of the blood is restored to normal and the blood therefore becomes relatively poor in corpuscles. In a few weeks, however, the corpuscular content of the blood is found to be once more normal, showing that the loss of corpuscles has been followed by a compensatory regeneration. The fact that the pigments constantly leaving the body with the urine and fæces, namely, urobilin and stercobilin, are derived from bilirubin, which arises from hæmoglobin, indicates that constant destruction of red blood cells must be proceeding. Since the number of corpuscles remains unaltered, this loss must be made good by a continual regeneration of fresh red corpuscles.

That the marrow is involved in the process is shown by the fact that it is the only tissue of the body which undergoes an alteration in appearance when blood formation is reduced by starvation, or stimulated by such means as repeated bleeding or by destruction of corpuscles by the injection of toxic agents. In adult animals there may be three forms of bone marrow, yellow, gelatinous and red. By starvation, especially in birds, the amount of yellow marrow is increased, at the expense of the red; when the animal is fed again the yellow marrow is changed first to the gelatinous and then to the red form. In the yellow marrow there is an abundance of fat cells, and the blood-forming tissue is small in amount, and dormant. In adult mammals, the yellow marrow normally occupies the shafts of the long bones, and the red marrow the cancellous spaces in such bones as the vertebræ, ribs and epiphyses. After extensive blood destruction, the red marrow is found to have increased in extent and in many cases to have invaded the greater part of the shaft of the bone, taking the place of the yellow marrow. It is in the red marrow, therefore, that we must seek the precursors of the red blood corpuscles.

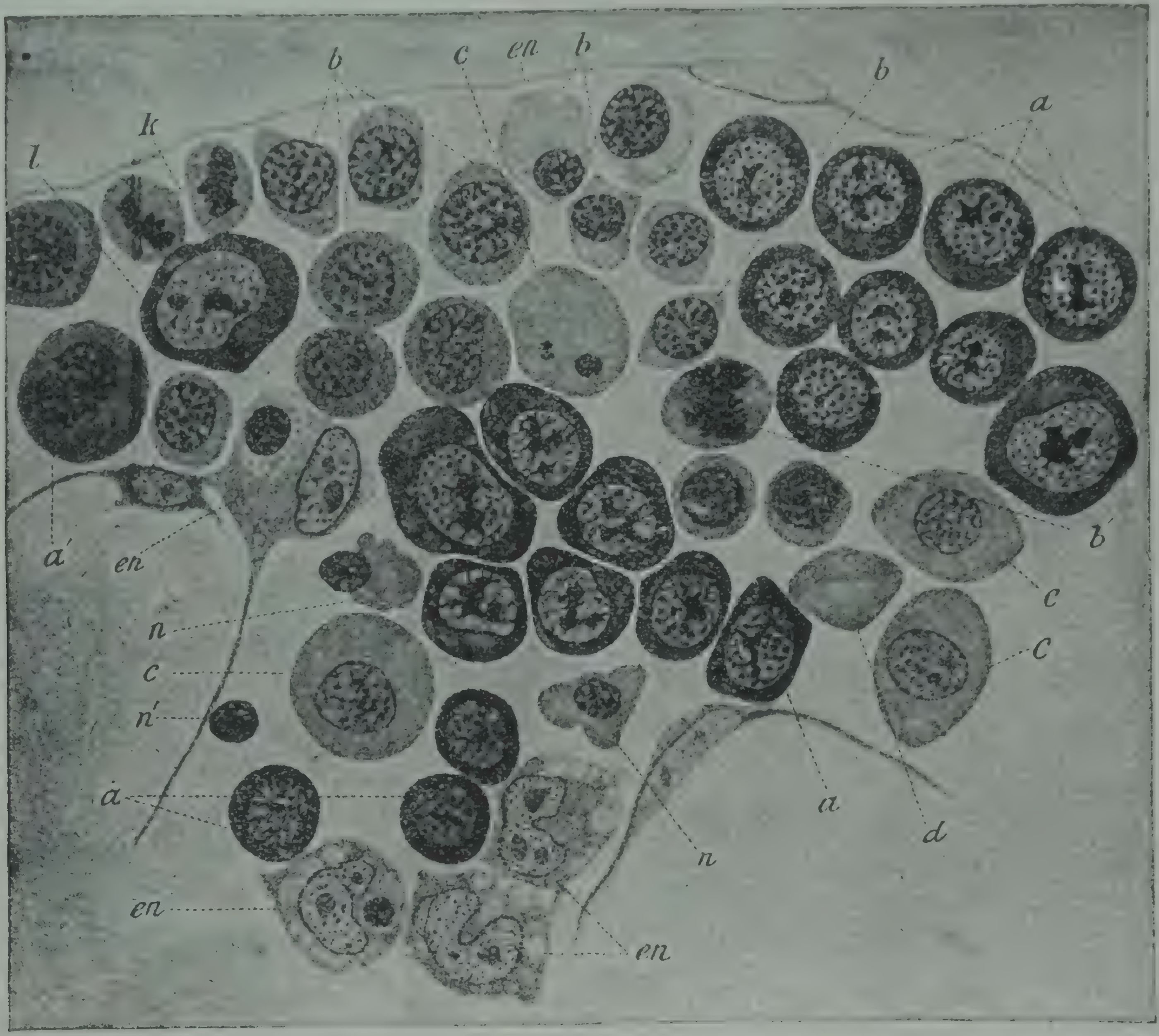


FIG. 333. Part of a blood vessel from the yolk sac of a rabbit embryo, showing the changes which occur in the formation of erythrocytes. (From SCHAFER after MAXIMOW.)

a. Megaloblasts.
b. Normoblasts, changing into erythrocytes.
c. Erythroblasts, in which the nuclei are dividing.
d. An erythrocyte fully formed, but not yet disc-shaped.
en. Phagocytic endothelial cells.

The blood vessels in the bone marrow are quite entire,* *i.e.* completely lined by endothelium, and consist, in addition to arteries and veins, of a very rich plexus of sinusoids. But it is a fact of great importance that not all of these sinusoids are open at any one time, many of them being collapsed and admitting no passage of corpuscles through them. This is particularly evident in the gelatinous marrow.

---

* DRINKER, DRINKER and LUND. *Amer. J. Physiol.*, 1922, **62**, 1.

The changes concerned in red-cell formation are most easily studied in birds in which the process has been checked by starvation. If such an animal is fed, it is found that many of the sinusoids become shut off as soon as blood formation commences, and it is in these closed vessels that the process of formation of the red cells occurs (Fig. 334); as soon as it is complete, the sinusoids open up, and the new cells are swept into the circulation.

The red cells are formed from the endothelial cells of the closed sinusoid; these swell and divide and become so-called *megaloblasts*. These then become *erythroblasts* and then *reticulocytes* by change and loss of the nucleus, and end by becoming *erythrocytes*. In the process of maturation, hæmoglobin is formed, probably at the expense of the nucleus, which, however, is still retained in the erythrocytes of ovipara.

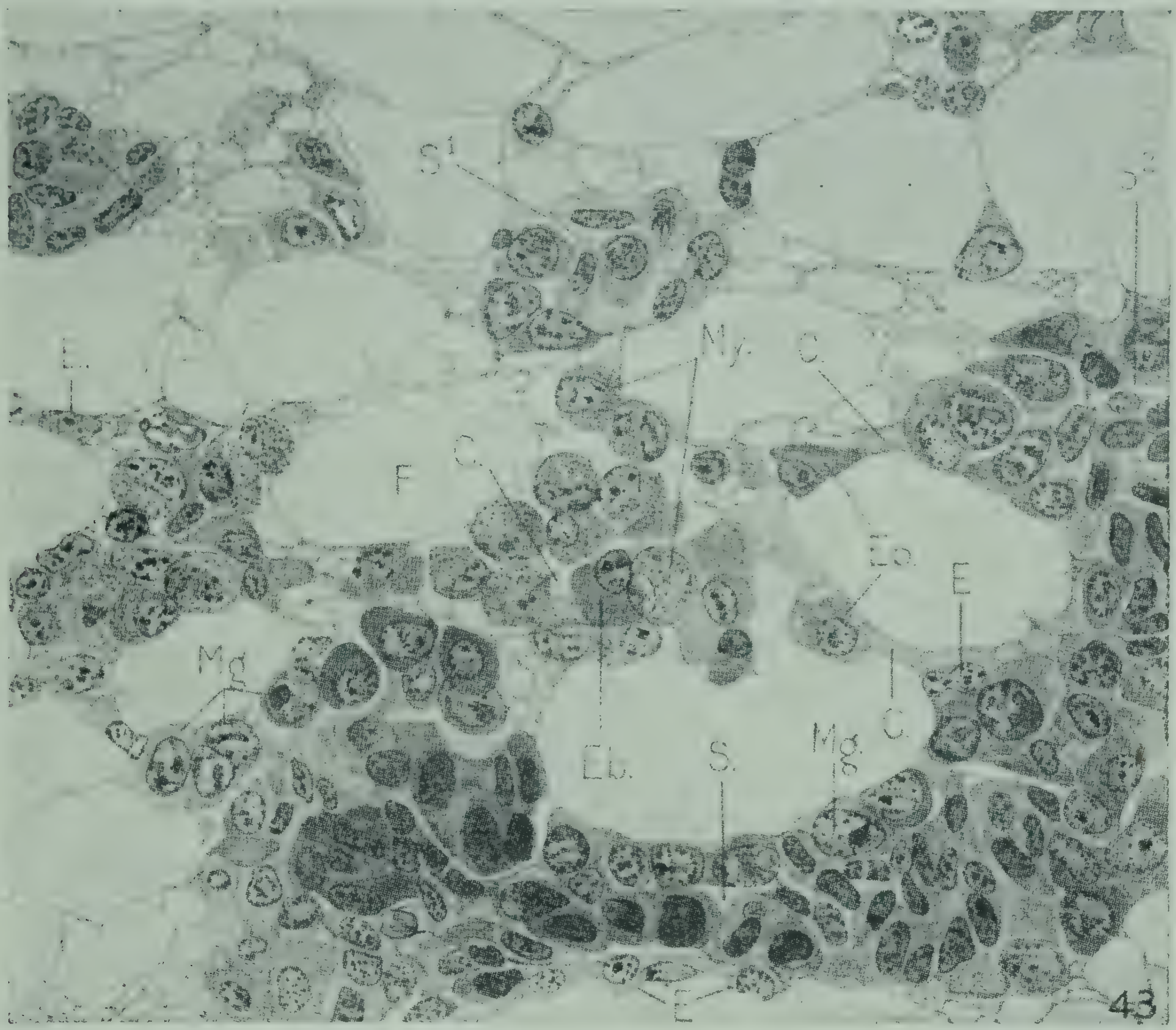


FIG. 334. Section of bone marrow of pigeon showing regeneration of red cells. Three sinuses, S, S^1, S^2, are seen filled with developing red cells. Along the inner wall are cells of the earlier phases—megaloblasts and early erythroblasts: note the difference in the depth of shade of cytoplasm as the hæmoglobin develops; in the centre of the sinuses are the older erythroblasts. In the middle of the field is a group of young myelocytes, *outside* a capillary, within which is a single erythroblast. Towards the right this capillary branches into two vessels, each containing a single erythroblast: one branch joins the sinus S^1, and the other joins sinus S^2.

C, capillary; E, endothelial cell; Eb, erythroblast; F, fat-cell; Mg, megaloblast; My, myelocytes; S, S^1, S^2, sinuses. (DOAN, CUNNINGHAM and SABIN. *Contrib.* to *Embryol.*, 1925, **83**, 165.)

In mammals the process is very similar, except that *megaloblasts* are not normally seen. The *erythroblasts* are often to be seen in process of division, and the nucleated daughter-cells appear to undergo a process of karyolysis, the nucleus slowly disappearing, sometimes with preliminary fragmentation, as the hæmoglobin increases in the developing cell. When blood formation is quickened as the result of previous destruction or loss, some immature and still nucleated blood discs may make their way into the circulation and be found in the blood, where they are spoken of as *normoblasts*. Other young, though non-nucleated, cells (*reticulocytes*) can also be distinguished by the fact that they stain with cresyl blue, while the older erythrocytes do not. In the blood of normal adults only about 1 per cent. of the red cells are reticulocytes. It takes about seven days for the passage from megaloblast to reticulocyte, and a further two days for conversion to the erythrocyte.

Examination of the bone marrow in the human subject (marrow biopsy) is carried out by puncture of the sternum and making a smear preparation of the extracted marrow.[*]

In various pathological states erythrocyte formation may, as in the fœtus, occur in liver, spleen or lymphatic tissue, as well as in the marrow: this is called extramedullary blood production.

Conditions for Red Cell Formation. The fact that the rate of formation of new red cells keeps pace with the rate of their destruction, indicates that there must be some stimulus which acts appropriately on the red marrow. The stimulus is more effective in the presence, or else is supplemented by the prevalence, of low oxygen pressures and *vice versâ*. Hence, the blood of animals exposed to atmospheres deficient in oxygen soon contains more corpuscles than when the animals breathe ordinary air, while in animals breathing oxygen-rich air the corpuscles are reduced in number.[†] Alternatively, increase of CO_2 in the blood, or of H-ion concentration, may be the effective stimulus.

Some factor is also necessary to promote maturation. We do not know its exact nature, but it is certainly in part of a chemical nature and emanates from the liver. Minot and Murphy[‡] discovered that feeding with raw liver has a most beneficial effect in pernicious anæmia, in which, in spite of great hypertrophy of the red marrow, conversion of megaloblasts to erythroblasts is deficient, and extracts of liver having the same effect can be prepared.[§] The administration is followed in four days by a sudden increase in the reticulocytes, which reach a maximum at about the eighth day. The active compound is called the *hæmopoietic factor*. A similar substance *ventriculin* has been prepared from hog's stomach, and there are indications that the substance in fact always originates in the stomach from which it is carried to, and stored in, the liver. The origin of the substance in the stomach has been supposed to be from a thermostable *extrinsic factor* soluble in 80 per cent. alcohol, present in certain foods, *e.g.* meat, yeast, and a thermolabile *intrinsic factor* (hæmopoietin), possibly an enzyme, formed in the stomach itself.[||] It has been suggested that folic acid, or a derivative of it, may be the extrinsic factor (Ch. XLII), but the nature of the factor is a complex question not yet capable of a final answer.[¶] In pernicious anæmia the intrinsic factor is lacking. It is a significant fact that pernicious anæmia is always associated with diminished gastric juice formation, achlorhydria and atrophic gastric mucosa.

The problem of the hæmopoietic factor was brought to a fresh focus by the extraction from proteolysed liver of a deep red substance having prodigious activity,[**][††] the maintenance dose being of the order of 1 to 2 μg daily. The compound has a molecular weight of about 1,500 and contains 4 per cent. of cobalt[‡‡]—an element previously shown to be essential for normal hæmopoiesis in some species.

Iron is, of course, essential for new red cell formation, and an iron protein complex, *ferritin*, plays a part in the absorption, transport and storage of iron.[§§] In rats this is more efficiently carried out if traces of copper are also present in the diet.[||||] The copper accelerates the conversion of assimilated iron into hæmoglobin.* Copper is present in the blood in organic combination with protein.† Riboflavin also accelerates hæmoglobin production: the protoporphyrin fraction of hæmoglobin is largely derived from glycine and acetic acid‡, and so can be synthesised in the body.

---

[*] Osgood and Seaman. *Physiol. Rev.*, 1944, **24**, 46.

[†] Campbell. *J. Physiol.*, 1926, **62**, 211.

[‡] Minot and Murphy. *J. Amer. Med. Assoc.*, 1926, **87**, 470.

[§] Cohn, Minot, *et al.* *Journ. Biol. Chem.*, 1928, **77**, 325.

[||] Castle. *Harvey Lects.*, 1934–5, **30**, 37.

[¶] Welch. *Feder. Proc.*, 1947, **6**, 471.

[**] Lester Smith. *Nature*, 1948, **161**, 638.

[††] Rickes, Brink, Konivszy, Wood and Folkers. *Science*, 1948, **107**, 396.

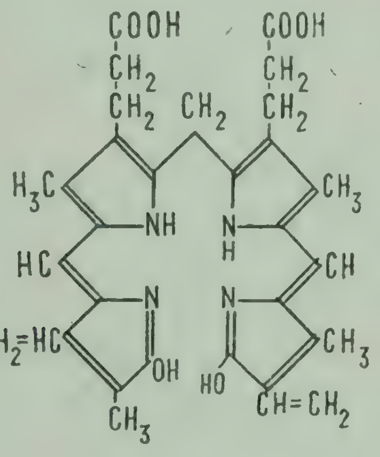
[‡‡] Lester Smith. *Nature*, 1948, **162**, 144.

[§§] Granick, *J. Biol. Chem.*, 1946, **164**, 737.

[||||] Hart *et al.* *J. Biol. Chem.*, 1928, **77**, 797; 1929, **84**, 115; Orton, Underhill and Lewis, *J. Biol. Chem.*, 1932, **96**, 1; Sheldon, *Brit. Med. J.*, 1932, 869.

Destruction of Red Cells. How long a corpuscle continues to exist in the circulating blood has been variously estimated, but feeding experiments with glycine labelled with N^{15} point to a life-span of about 127 days.‡ Sooner or later, however, every corpuscle undergoes disintegration, a process generally ushered in by the ingestion of the corpuscle by some phagocyte cell. Thus, in the hæmolymph glands and in the spleen, we find large reticulo-endothelial cells which have englobed red corpuscles and in which we can recognise pigment granules derived from their destruction; this process of destruction of red cells by macrophage cells has been seen to occur in tissue cultures under the microscope.§ The chief places of disintegration are certainly the liver and spleen, and in these organs the hæmoglobin is converted first into the bile pigment biliverdin and then into bilirubin, which is then excreted by the liver. The liver cells proper do not form bilirubin, but merely excrete the bilirubin which is formed in the various cells of the reticulo-endothelial system. Any methæmoglobin, however, is not so converted, but may be in part changed to coproporphyrin III.

A probable structural formula for bilirubin is:—



Bilirubin

In addition to removal of the globin from Hb, it will be seen, by comparison with the formula of hæm (p. 516), that the ring has been opened by oxidation at the α-position, the CH group at γ reduced, and, most important of all, the iron removed from within the molecule. These changes are probably effected in stages || and can be demonstrated *in vitro*. The opening-up probably comes first and may occur while the ferroporphyrin part is still attached as a prosthetic group to the globin, or to a denatured globin. It results from the coupled oxidation, by free oxygen, of hæmoglobin or (more easily) of hæmochromogen and of oxidisable substances, such as ascorbic acid or glutathione, and leads to the formation of a variety of green compounds, known as choleglobins, cholehæmochromogens, verdohæmochromogens, etc.

Since hæmochromogens are most easily oxidised in this way, yielding cholehæmochromogens, it has been suggested that one function of the reticulo-endothelial cells is to denature the globin part of the Hb molecule, so forming hæmochromogens, which then are oxidised to cholehæmochromogens. Verdohæmochromogen is a simpler analogue derived from pyridinehæmochromogen.¶

From compounds of the verdohæmochromogen type the iron is readily

---

[*] Elvehjem and Sherman. *J. Biol. Chem.*, 1932, **98,** 309.

[†] Mann and Keilin. *Proc. R.S.*, 1938, **126B**, 303.

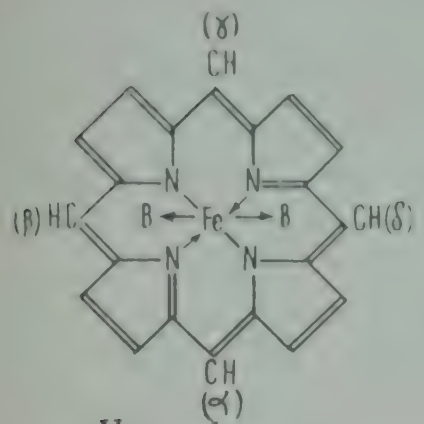
[‡] Shemin and Rittenberg. *J. Biol. Chem.*, 1945, **166,** 627.

[§] Rich. *Bull. Johns Hopkins Hosp.*, 1924, **35.** 415.

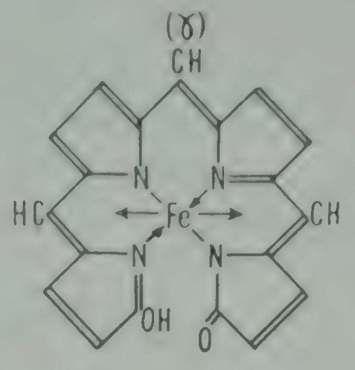
[||] Lemberg. Perspectives in Biochemistry. (Cambridge 1937, p. 137.)

[¶] Lemberg. *Biochem. J.*, 1935, **29,** 1322; 1941, **35,** 363: *J. Proc. Austral. Chem. Inst.*, 1939, **6,** 170.

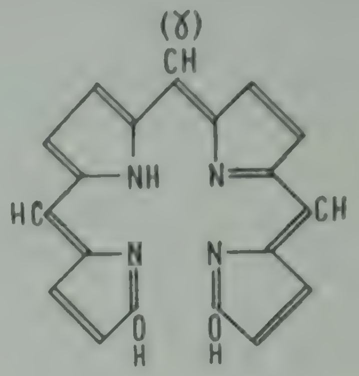
removed by the action of acids, with formation of biliverdin, and this by reduction of the γ group is transformed into bilirubin. The relations are shown below, in skeleton form. (B = base.)


(γ)
CH
(β) HC
CH (δ)
B
Fe
B
N
CH
(α)

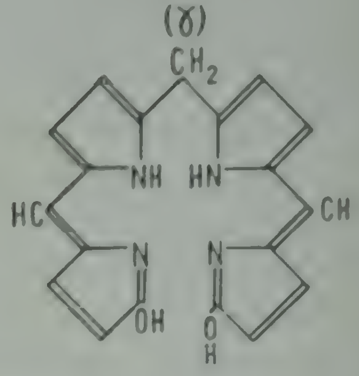
Hæmochromogen


(γ)
CH
HC
CH
Fe
N
OH
O

Verdohæmochromogen

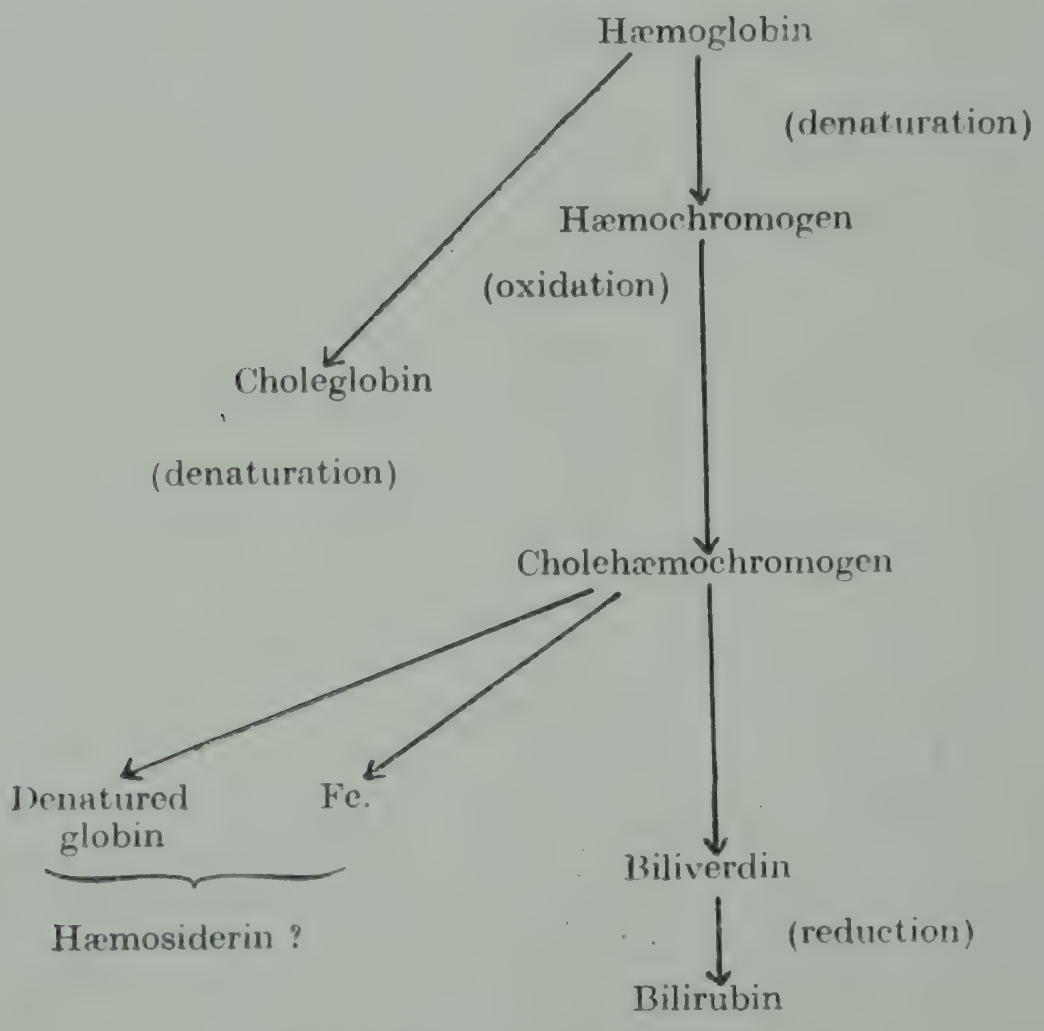

(γ)
CH
HC
CH
NH
N
O
H

Biliverdin


(γ)
CH₂
HC
CH
NH
HN
N
OH
O
H

Bilirubin

Quite possibly the iron and the denatured protein are removed together and in loose combination. These changes can be depicted schematically :—


Hæmoglobin
(denaturation)
Hæmochromogen
(oxidation)
Choleglobin
(denaturation)
Cholehæmochromogen
Denatured globin
Fe.
Hæmosiderin ?
Biliverdin
(reduction)
Bilirubin

The iron remains in combination in the cells of the liver, spleen, etc., in the form of brown granules of a substance called *hæmosiderin*. This contains about 17 per cent. of iron, probably as a colloidal ferric oxide adsorbed on an organic substratum.* It is probably re-utilised for the formation of fresh hæmoglobin in the bone marrow. A section of normal liver immersed in potassium ferrocyanide and then in acid alcohol shows the presence of iron by blue granules in the liver cells. The amount of iron which can be demonstrated in the liver in this way is enormously increased in any condition which augments the rate of blood destruction, *e.g.* in pernicious anæmia or by the action of poisons which cause extensive hæmolysis.

Reduction of biliverdin to bilirubin is probably effected by enzymes, *e.g.* lactic dehydrogenase, and is correlated with carbohydrate metabolism ; in individuals or species with low liver glycogen biliverdin is excreted as such.

The formation of bile pigment *via* cholehæmochromogens may possibly occur under certain conditions in the blood itself, though its principal seat is in the cells of the reticulo-endothelial system.

---

* Cook. *J. Biol. Chem.*, 1929, **82**, 595.

That the liver is by no means the sole seat of bile pigment formation is shown by the results of experiments by Whipple, F. C. Mann and their colleagues. The latter find that the liver can be completely removed from dogs, provided sufficient glucose is given to prevent the occurrence of hypoglycæmic symptoms ; ten hours or less after removal of the liver definite jaundice develops, and the urine, blood plasma and adipose tissue are stained with bilirubin. The jaundice is increased if hæmolysed blood is injected.

When hæmoglobin is set free into the circulation by hæmolysis, or is injected, it is removed in three ways ; by excretion in the urine, by absorption by reticulo-endothelial cells, to form bilirubin and hæmosiderin, and by breakdown in the blood stream forming methæmalbumin, an analogue of methæmoglobin.* This is eliminated by the liver and finds its way into the fæces as coproporphyrin III.

Bilirubin is present in small amount (0·15 mg./100 c.c.) in normal blood plasma in some species, including man, but in jaundice the amount is largely increased. Its presence can be demonstrated spectroscopically, or by the *van den Bergh reaction.* In this test, the serum is precipitated with alcohol and the filtrate added to a freshly prepared mixture of sulphanilic acid, hydrochloric acid and sodium nitrite, when a violet-red colour is produced.

For Reference

Isaacs (1937). Formation and Destruction of Red Blood Cells. *Physiol. Rev.*, **17**, 291.
Jordan (1942). Extramedullary Blood Production. *Physiol. Rev.*, **22**, 375.
Maximow (1924). Relation of Blood Cells to Connective Tissues and Endothelium. *Physiol. Rev.*, **4**, 533.
Michaelis (1947). Ferritin and Apoferritin. *Adv. in Protein Chemistry*, Vol. III, p. 53. (Acad. Press Inc., N.Y.)
Rich (1925). The Formation of Bile Pigment. *Physiol. Rev.*, **5**, 182.
Robscheit-Robbins (1929). The Regeneration of Hæmoglobin and Erythrocytes. *Physiol. Rev.*, **9**, 666.
Rous (1923). Destruction of the Red Corpuscles. *Physiol. Rev.*, **3**, 75.
Sabin (1922). On the Origin of the Cells of the Blood. *Physiol. Rev.*, **2**, 38.
Sabin (1928). Bone Marrow. *Physiol. Rev.*, **8**, 191.
Schultze (1940). Metallic Elements and Blood Formation. *Physiol. Rev.*, **20**, 37.
Watson and Larson (1947). The Urinary Coproporphyrins in Health and Disease. *Physiol. Rev.*, **27**, 478.

THE WHITE BLOOD CORPUSCLES

These number usually about 9,000/c.mm., and are of various kinds, some of which exhibit amœboid movements. Ehrlich showed that some corpuscles contained granules which were stained by acid dyes such as eosin, and these corpuscles he called *eosinophile.* Others, the *basophile* cells, were stainable with basic dyes, while a third type were only lightly stainable, and were called *neutrophile* leucocytes. The white corpuscles are classified as shown (Fig. 335), the figures expressing the average results of what is called a *differential cell count.* This shows, out of a total of 100 white cells :

(1) Granulocytes.
  (*a*) Neutrophile. Approximate average, 68 to 70 per cent.
  (*b*) Eosinophile. Average, 3 per cent.
  (*c*) Basophile (mast cells). Average, 0·5 per cent.

(2) Monocytes.
  (*a*) Large mononuclears. } Average, 4 per cent.
  (*b*) Transitional. }

(3) Lymphocytes. Average, 23 to 25 per cent.

The various white cells have the following characters :

(1) Granulocytes. (*a*) *The Neutrophile Cells,* or *polymorphonuclear leucocytes.* These present a lobed nucleus and their protoplasm contains abundant fine neutrophile granules

---

* Fairley. *B.M.J.*, 1940, *ii.*, 213.

which stain with Ehrlich's triacid and with the May-Giemsa stains. The cytoplasm contains proteolytic enzymes capable of causing autolysis (hence the liquefaction of pus on standing). These cells are actively amœboid and phagocytic, and their numbers rapidly increase in pyogenic infections, in which they form pus cells.

(*b*) *The Eosinophile Cells.* The nucleus is generally single, but is often crescent-shaped or reniform. The protoplasm is crammed with large, discrete, highly refractive granules, which stain deeply with eosin and give micro-chemical reactions for iron as well as phosphorus. A large fraction of the histamine of the blood is present in these cells.* The eosinophile cells are not normally phagocytic.

(*c*) *The Basophile* (*or Mast*) *Cells.* These, which are somewhat smaller than the polymorphonuclear cells, have a lobed or tri-lobed nucleus and present in their protoplasm a number of granules which stain deeply with basic dyes. In normal blood the basophile cells are rare, though they appear, sometimes in large numbers, in disease, and are normally present in the blood of some of the lower animals.

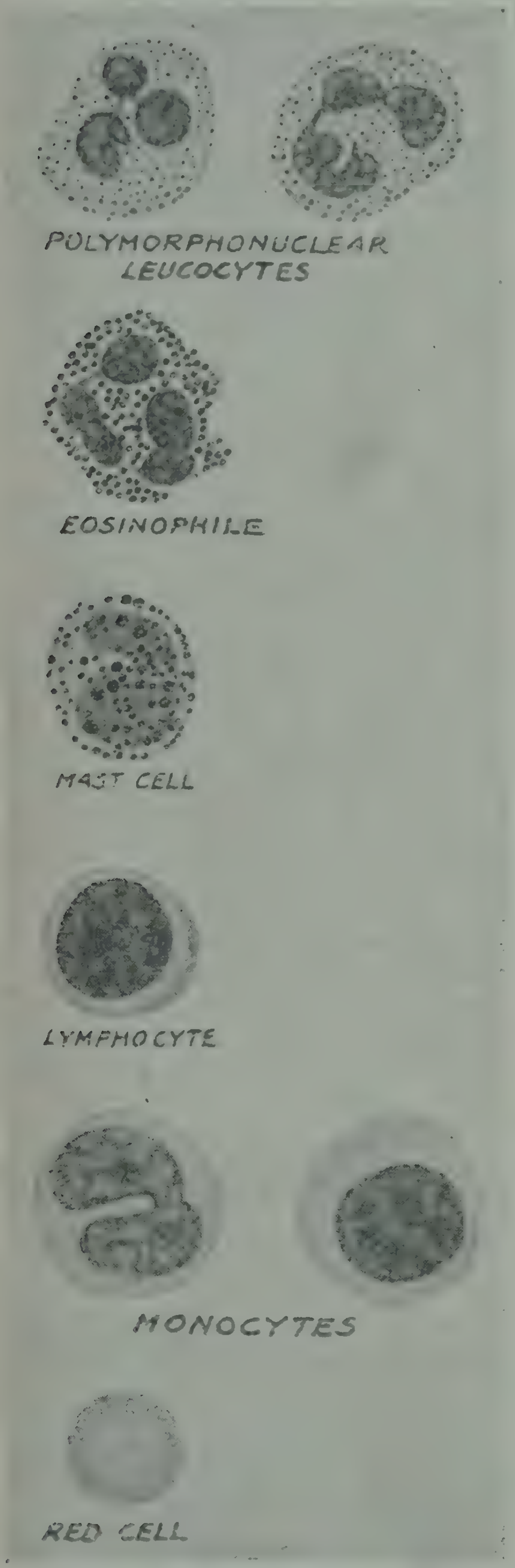


Fig. 335. Various forms of Blood Corpuscles.

(2) Monocytes. These are large vacuolated cells with a single nucleus, and with fairly abundant faintly granular cytoplasm. The large mononuclears are rare cells in normal blood, but they are of considerable importance. They have the property of ingesting foreign particles of almost any sort, and for this reason are able to take up certain intravitam dyes, such as pyrrol or trypan blue, or granules of Indian ink or carmine.

The *transitional cells* are more numerous, as a rule, than the large mononuclears, forming about 4 per cent. of the total whites. Their general structure resembles that of the monocytes, but they are smaller. Their nucleus is indented or slightly lobed, and the cytoplasm has fine granules which stain with azur.

(3) Lymphocytes. These abundant cells, smaller than the other mononucleated white cells, and of more uniform diameter (about 8 μ), have a round nucleus and only a small amount of cytoplasm, which has very few granules, but such as it has are coarse and stain with azur.

The total white count is high in infants and during pregnancy.

The errors incidental to both total white and differential counts are considerable, but after due allowance has been made for them, surprisingly large and rapid fluctuations in the leucocyte count and differential are demonstrable. Such rapid alterations are due, not to altered rates of formation and destruction, but to altered distribution within the circulation. In the smallest veins, and in wide capillaries, neutrophiles dawdle in the peripheral stream, particularly when the flow is slowed down, and, similarly, lymphocytes tend to lag in the lymphatic system. Any circumstance which alters the circulatory conditions is therefore liable to accentuate or diminish these tendencies; hence, exercise causes an increase in the total count, with an altered differential picture: if carried to the point of fatigue the total count may show a 300 per cent. increase on the resting figure.

Formation of the Leucocytes. The white cells are continually being lost and replaced. The normal life of a leucocyte is variously estimated at four days or upwards, but many die prematurely. The question as to the origin of each kind of leucocyte cannot be regarded as settled. According to Maximow and others, all the cells, red and white, are derived from one kind of cell called a *hæmocytoblast.* This opinion seems now to be losing ground, in favour of the view of Sabin, according to which erythrocytes and macrophages are of endothelial origin, while the granulocytes, lymphocytes and monocytes arise from a common parent cell. This parent, or *reticular cell* is converted to a *primitive white cell*, examples of which are found in lymph glands, spleen and bone marrow. The subsequent history of these cells differs according to their situation, the two chief sites being lymph gland and marrow.

---

* Code. *J. Physiol.*, 1937, **90**, 349, 485.

Lymphocytes are derived from the reticular cells of the lymphatic glands and the lymph nodules surrounding mucous cavities. These lymphatic nodules present towards their centre a clearer zone, consisting of lymphoblasts. By the division of these cells lymphocytes are formed, pushing towards the periphery of the nodule, where they make their way into the lymph sinus and are carried in large numbers (200 million hourly in the dog) by the lymph into the blood. Where they then go, or what their fate may be, is unknown.

In the red bone marrow granulocytes originate, outside the blood vessels, from the primitive white cells, which may be seen to be formed in abundance from the reticular cells during the gelatinous phase of the marrow. From these, *myeloblasts*, and then *myelocytes* of various types are formed, and divide actively, forming groups of cells. These contain a single rounded nucleus and a well-marked protoplasm which may contain granules, generally eosinophile in character, but sometimes basophile. All intermediate states are to be found in the bone marrow between these myelocytes and the polymorphonuclear and eosinophile leucocytes. When mature, the granulocytes pass through the walls of the sinusoids into the blood stream. In leukæmia there may be myeloblasts in the blood as well as an increase either of eosinophile cells or of neutrophile cells, and either condition is associated with changes in the red bone marrow.

The monocytes are principally formed from reticular cells in the lymph glands and spleen, though exceptionally they, and also lymphocytes, may arise in the marrow. In certain abnormal states, clasmatocytes may also be present in the blood. We may therefore provisionally arrange the leucocytes of the blood according to their origin as follows :

(1) Small lymphocytes derived from lymphoid tissue.

(2) Monocytes, formed in the lymph glands and spleen, or exceptionally from reticulo-endothelial cells in other situations.

(3) Polymorphonuclear leucocytes formed in bone marrow.

(4) Eosinophile cells derived from similar cells in the bone marrow.

In certain diseases, the primitive white cells may develop along lines which are abnormal for the situation in which they occur ; thus myeloblasts may appear in the spleen, or even in lymph glands, and monocytes in the marrow.

Functions of the Leucocytes. The polymorphonuclear leucocytes present two phenomena, that of amœboid movement and that of ingesting foreign particles by a process known as "phagocytosis." Many unicellular organisms and the undifferentiated cells of many kinds of tissue behave similarly. All the phenomena connected with the process of *inflammation* in higher animals are directed to the assemblage of polymorphonuclear leucocytes at the spot which is the seat of injury or of infection, so that they may devour and remove either the injured tissue or the invading micro-organisms. This process therefore plays an important part in determining the immunity of any animal against infection.

The Reticulo-endothelial System. In addition to the polymorphonuclear leucocytes, certain other cells have phagocytic properties. These are the monocytes of the blood, and certain widely dispersed tissue cells. All of these arise, like the blood cells, from the mesenchyme of the embryo. Of the mesenchyme cells, some become the *fibroblasts* of the connective tissues. A number, however, retain their embryonic developmental potentialities, and become converted into tissue histiocytes belonging to one of two types, called respectively the *free and the fixed histiocytes.* The *free histiocytes,* wandering cells, or macrophages, have phagocytic properties, and are to be found in the connective tissues. By a change in their granulation they may become the *mast cells,* familiar as inhabitants of the connective tissues. The *fixed histiocytes,* or "resting wandering" cells, are to be found in large numbers in the serous membranes, especially in the omentum, and also as histiocytes which have taken up the position of an endothelial lining. This scattered system of specialised "endothelial" cells is sometimes called the *reticulo-endothelial system.* Examples are the so-called "endothelial cells" which form the incomplete lining of the sinusoids of the liver and spleen. In the *liver* the cells are known as the *Kupffer* or stellate cells. This widely dispersed system comprises cells of very differing function, shape and appearance, but they all possses in common the property of ingesting foreign particles of varying grades of size. Even relatively coarsely particulate matter is taken up by these cells, *e.g.* by the Kupffer cells of the liver. Thus, if Indian ink or carmine is injected intravenously, granules of the pigment are to be found shortly afterwards in the Kupffer cells, and in the similar cells of the spleen and marrow, as well as in the polymorphonuclear leucocytes and monocytes of the blood stream. All foreign colloidal matter, including colloidal dyes, *e.g.* pyrrol blue, is also taken up by these cells.

Whenever any surplus or dead tissue has to be cleared away, whether as the result of injury or in the course of the metamorphosis of organs, *e.g.* the tail of the tadpole, the phagocytic cells of the tissues play an important part. In mammals, including man, the moulding of the long bones which occurs in the process of growth is effected by continual and coincidental processes of absorption and new formation of bone. The absorption is carried out by means of special phagocytes called osteoclasts.

The other functions, which have been ascribed to leucocytes, are all of them questionable. Thus some authors ascribe to leucocytes a significant part in the taking up of fat from the intestine and its carriage into the lymphatic system. In the coagulation of the blood, the leucocytes have been supposed to act by the discharge of substances which may be precursors of thrombin. The leucocytes and lymphocytes are also believed to play a part in the formation of anti-toxins.

For Reference

BUNTING (1922). The Leucocytes. *Physiol. Rev.*, **2,** 505.

CHÈVREMONT (1948). Le système histiocytaire ou réticulo-endothélial. *Biol. Rev.*, **23,** 267.

GARREY and BROWN (1935). Variations in White Blood Cell Counts. *Physiol. Rev.*, **15,** 597.

SABIN (1922). On the Origin of the Cells of the Blood. *Physiol. Rev.*, **2,** 38.

SABIN (1928). Bone Marrow. *Physiol. Rev.*, **8,** 191.

SACHS (1926). The Reticulo-endothelial System. *Physiol. Rev.*, **6,** 504.

STURGIS and BETHELL (1943). Quantitative and Qualitative Variations in Normal Leukocytes. *Physiol. Rev.*, **23,** 279.

THE BLOOD PLATELETS

In addition to red and white cells, appropriately fixed and stained blood films also present a number of granular bodies, most of which are about 2μ in diameter, and which tend to aggregate together in clumps. These are the platelets or thrombocytes. Their number usually ranges from 280,000 to 760,000 per cubic millimetre. Their shapes vary considerably. Their central part is more strongly refracting than the periphery and stains with basic dyes, but is not regarded as a nucleus (Figs. 336, 337).

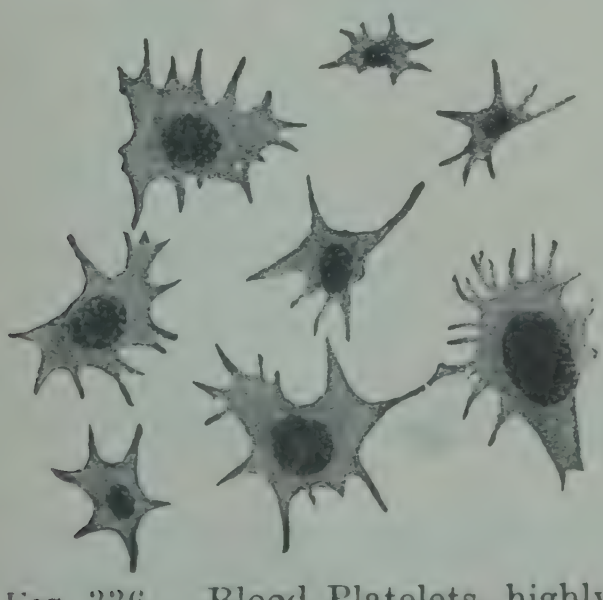


FIG. 336. Blood Platelets, highly magnified, showing the amœboid forms which they may assume. (After KOPSCH.)

Similar platelets are observed when the blood is received into normal salt solution; as the mixture clots, the filaments of fibrin can often be seen to radiate from a clump of disintegrated blood platelets as from a centre. They are enumerated by pricking the finger through a drop of 2·5 per cent. sodium citrate solution in normal saline. A drop is transferred, after further dilution with citrate-saline in marked vessels, to a slide, and the platelets and red cells counted in appropriate fields. From the ratio between the two, and a separate red cell determination, the number per cubic millimetre are calculated. If the blood of an animal be defibrinated and returned to the circulation, it will be found for the next few days to be quite free from blood platelets. Adrenaline injections raise the platelet count, and so does splenectomy, from which facts it may be inferred that depôts of platelets are lodged in the spleen. Most histologists regard the platelets as derived by fragmentation from the cytoplasm of the giant cells, or *megakaryocytes*, of the bone marrow or, as some say, of the lungs. Injection of pyridine appears to release showers of new and sticky platelets into the blood stream.*

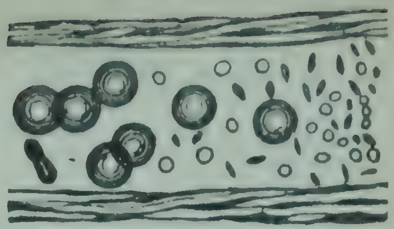


FIG. 337. Blood Corpuscles and Blood Platelets within a small Vein. (SCHAFER after OSLER.)

THE COAGULATION OF THE BLOOD

The subject of blood coagulation presents so many aspects on which there is disagreement that it will be necessary here to deal chiefly with what appear to be facts, and, since no other course is consistent both with brevity and with clarity, to treat the theory in a brief and dogmatic manner. The particular theory which will, for this purpose, be upheld, is that originally due to Morawitz.†

In the process of coagulation, while the red and white corpuscles remain almost or quite intact, the plasma becomes solid from the production in it of a network of fibrin. The coagulation takes place most rapidly at body temperature and is greatly slowed by cold. There are many patterns of coagulometer by which the clotting time can be measured,‡ or graphically recorded.§ Under constant conditions the time is constant for any individual. Under the ultramicroscope, the fibrin is seen to have a crystalline structure.‖ The question of coagulation involves the consideration of the precursor of fibrin and of the conditions which determine its conversion into fibrin.

It will be clearer if the theory is outlined to begin with. The factors in the blood which are concerned with clotting are, fibrinogen, prothrombin, platelets and calcium ions. Briefly, the precursor of fibrin which is present in the fresh plasma is the protein *fibrinogen*. It is converted into fibrin through the agency of an enzyme-like substance called *thrombin*, which is formed in the shed blood from a precursor called *prothrombin ;* prothrombin exists in the normal plasma. The prothrombin is converted into thrombin by an activator, *thromboplastin* (or *thrombokinase*) ; thromboplastin alone only slowly activates prothrombin, but in presence of calcium salts the activation is rapid. Thromboplastin is not present in circulating blood, but is formed from many kinds of tissue when these are damaged, and, among others, from the blood platelets. Hence, normal blood contains all the factors for clotting—platelets which form thromboplastin ; and prothrombin, which this activates in presence of the calcium salts, thus forming thrombin. The thrombin converts the fibrinogen into fibrin.

According to this theory, then, the prime mover in the series of changes terminating in coagulation is thromboplastin. We may now consider some of the facts.

Methods of Preventing Coagulation. It is evidently impossible directly to arrive at any conclusions during the few minutes which elapse before the appearance of the clot. By retarding coagulation we can obtain a cell-free plasma by centrifugalisation, after which the clear supernatant plasma can be drawn off and its coagulation studied.

(1) Normal blood in contact with the normal uninjured vessel intima remains fluid. If a vein be tied in two places and excised, the blood contained between the ligatures will remain fluid, sometimes for days. On bringing it into contact with foreign substances, it undergoes coagulation.

---

* WRIGHT. *J. Path. and Bact.*, 1944, **61**, 151.

† MORAWITZ. *Arch. klin. Med.*, 1904, **79**, 1, 25, 215. *Hofmeister's Beitr.*, 1904, **4**, 381.

‡ ADDIS. *Q. J. Exp. Physiol.*, 1908, **7**, 314 ; GIBBS. *Q. J. Med.*, 1924, **17**, 312.

§ STOKER. *J. Scientif. Inst.*, 1934, **11**, 315.

‖ HOWELL. *Amer. J. Physiol.*, 1914, **35**, 143.

(2) When an incision is made in the ordinary way into a blood vessel of a bird, the issuing blood clots very rapidly. But if the vessel be isolated and a perfectly clean glass cannula be inserted into it, blood drawn off into a clean sterilised beaker will remain unclotted for days.* It is very difficult to carry this out in the mammal.

(3) Clotting can be excited even in the living vein by introducing into the blood any foreign solid, *e.g.* a thread, which is wetted by the blood. If contact of shed blood with such substances be prevented by receiving it into vessels made of perspex, or of glass coated with a water-repellant layer, *e.g.*, paraffin, or with a monomolecular layer of *silicone*, and scrupulously free from dust, clotting may often be delayed for many hours.

(4) *Action of Cold.* If horses' blood, which always clots slowly, is cooled rapidly to between 0° C. and 1° C., it remains fluid for an indefinite time. The corpuscles sink, and the supernatant plasma can be decanted and even filtered.

(5) *Methods involving Mixture with Neutral Salts.* Blood is received into one-quarter its bulk of a 25 per cent. solution of magnesium sulphate, or into an equal volume of half-saturated sodium sulphate solution.

(6) *Methods depending on Decalcification of the Blood.* Oxalate plasma is obtained from blood received direct into a solution of sodium oxalate, so that the total blood contains 0·1 per cent. of the oxalate. Instead of oxalate we may use sodium fluoride (0·3 per cent.) or sodium citrate (0·35 per cent.).

(7) *Methods depending on the use of certain Substances of Animal or Plant Origin.* (*a*) *Peptone* plasma got from blood obtained by injecting rapidly into the veins of a dog or cat in a fasting condition a solution of commercial peptone in the dose of 0·1 g. per kg. of the animal. This causes a state of shock. On drawing blood immediately after the fall of blood-pressure has taken place, it is found to be incoagulable. The addition of peptone to blood *in vitro* does *not* inhibit clotting.

(*b*) *Leech Extract or Hirudin.* There is an anticoagulating substance in the buccal glands of the leech.† It will abolish the coagulability of the blood *in vitro*.

(*c*) *Heparin.* An anticoagulant substance, heparin, was first prepared from liver.‡ It is present also in muscle, lung, and various other tissues,§ and may be obtained as a powder, of which 1 mg. will prevent the coagulation, at 0° C., of from 100 to 500 c.c. of blood. At room temperature 1 mg. will inhibit coagulation of 10 to 20 c.c. of blood. A unit is 0·01 mg. of beef heparin.‖ (The activity varies with the species.) It acts equally well *in vivo* or *in vitro*. Chemically, heparin is a polysulphuric ester of mucoitin, a substance related to chondroitin. Mucoitin is a condensation compound of glucuronic acid and acetylated glucosamine.¶ Heparin is precipitated by protamine, and although either substance alone will prevent coagulation, the two together are ineffective.

(*d*) "*Dicoumarin*" present in spoiled sweet clover, when injected, or even when taken by mouth, causes reduction or loss of blood coagulability *after a few days*.

(8) *Methods depending on the use of Synthetic Organic Substances.* (*a*) *Chicago Blue*, 6B, injected intravenously or added to blood *in vitro* prevents coagulation.**

An even better, and less strongly coloured dye is *chlorazol fast pink*, B.K.S.††

(*b*) "Bayer 205" ("Germanin") ‡‡ or "Liquoid" (Roche), described as a sodium poly-anetholesulphonate,§§ injected intravenously also causes incoagulability.

Factors in Coagulation. By centrifugalisation of blood the coagulation of which has been delayed or prevented in any of these ways, we obtain blood plasma. The conditions which will bring about coagulation in such plasmata are strikingly diverse. Thus, in cooled plasma a rise in temperature is sufficient, though if the cooled plasma be filtered at 0° C. through a fine filter, it loses this spontaneous coagulability on warming. It can still be made to clot by the addition of certain substances, such as blood serum or the washings of a blood clot, and in some cases by the addition of tissue extracts. Oxalate plasma clots on addition of calcium salts.

Calcium Ions. We may take oxalate plasma as our type. It contains fibrinogen, prothrombin, a little thromboplastin, and an excess of oxalate. If calcium chloride be added in slight excess, and the plasma kept warm, it will clot. The place of calcium can be taken by strontium. Coagulation is also prevented by sodium fluoride, which precipitates calcium, and by sodium citrate, metaphosphate or pyrophosphate, which lead to the formation of a salt in which the calcium is in a complex anion. This shows that calcium must be present in the ionic form. The optimal $Ca^{\cdot\cdot}$ concentration is 5 mg./100 c.c., which is about that present in normal plasma, where about half the calcium is present in the ionic state. Calcium ions probably accelerate the activation of prothrombin by thromboplastin.*

---

* Delezenne. *Arch. de Physiol.*, 1897, **9** (Ser. 5), 333.

† Haycraft. *Arch. f. exp. Path.*, 1884, **18**, 209.

‡ Howell. *Bull. Johns Hopk. Hosp.*, 1928, **42**, 199.

§ Best. *B.M.J.*, 1938, *ii.*, 977.

‖ For standardisation, *v.* MacIntosh. *Bioch. J.*, 1941, **35**, 770, 776.

¶ Jorpes and Bergström. *Z. f. physiol. Chem.*, 1936, **244**, 253.

** Rous, Gilding and Smith. *J. Exp. Med.*, 1930, **51**, 807.

†† Huggett and Rowe. *J. Physiol.*, 1934, **80**, 82.

‡‡ Stuber and Lang. *Bioch. Ztschr.*, 1928, **194**, 204.

§§ Fleisch *et al.* *Pflüger's Arch.*, 1932, **230**, 814.

FIBRINOGEN. Plasma contains a protein, *fibrinogen*, not present in the serum, and which is the precursor of fibrin. Fibrinogen is a globulin precipitated by half-saturation with sodium chloride; it coagulates between 56° and 60° C., whereas serum does not give any precipitate at 60° C. For coagulation experiments it is best made from acidified oxalated plasma by freezing and slowly thawing out, by which it is precipitated; it is then collected in cold saline and dissolved by warming.† The solution of fibrinogen thus purified is perfectly clear and colourless.

THROMBIN. If calcium chloride be added to a purified solution of fibrinogen, no clotting occurs; but the purified fibrinogen can still be made to clot by the addition of fresh serum, due to the presence of thrombin in the serum. The last change in the act of clotting is therefore the change from fibrinogen to fibrin by the intervention of thrombin. It cannot be at this stage of the process that the calcium salts exercise their influence, since thrombin will cause the coagulation of oxalate plasma. Thrombin is a protein substance, is usually regarded as an enzyme (probably proteolytic), and is active in minimal quantities. When in watery solution it is readily destroyed at 60° C.; up to 50° C., rise of temperature hastens its action. The speed with which clotting occurs in fibrinogen solutions or oxalate plasma is directly proportional to the thrombin concentration; hence dilution retards clotting. After formation of fibrin much thrombin is adsorbed on it.

Thrombin may be provisionally defined as an enzyme capable of producing clotting in calcium-free fibrinogen solutions. Accepting this definition, we can show that thrombin is absent from circulating blood. If blood be received direct from the vessels into absolute alcohol and the precipitate, after coagulation by alcohol, be extracted by water, the extract is found to contain no trace of thrombin. Nor is thrombin present in fresh oxalate plasma. But this contains the precursor of thrombin, a protein which is called *prothrombin*, together with traces of another substance essential for the formation from it of thrombin, namely *thromboplastin*. Thrombin is best prepared from prothrombin by the action of thromboplastin in presence of calcium ions. Hence, addition of calcium salts causes oxalate plasma to clot, and the serum formed is found to contain thrombin. In the early stages of coagulation, therefore, there is an actual formation of thrombin. A *crude* solution of fibrinogen treated with calcium chloride also rapidly clots. This is because it also contains both prothrombin and thromboplastin. Our next step must be to inquire into the precursors of thrombin and the conditions of its formation.

PROTHROMBIN is present in normal blood plasma. Controversy has centred round the methods for estimating the prothrombin content of the blood: essentially Quick's method depends on determining the clotting time of recalcified oxalate plasma to which excess of thromboplastin has been added. For man the time is about 12–13 seconds.

In the two-stage method ‡ oxalated plasma is first defibrinated by adding a small amount of thrombin and removing the clot. The serum is then treated with calcium chloride and thromboplastin to convert all its prothrombin to thrombin. The concentration of the latter is then determined by finding the dilution of the serum needed to coagulate a standard fibrinogen solution in 15 seconds. The total final dilution of original plasma represents the number of " units " of prothrombin per c.c. (normal = 300–400 units/c.c.).

The results of the two methods sometimes differ seriously, and one reason advanced for the discrepancy is that another factor is necessary for rapid conversion of prothrombin to thrombin. Quick called this "prothrombin-A" a labile factor not absorbed by alumina. Owren * calls it factor 5, and studied a clinical case in which prothrombin time was prolonged because this factor was lacking: addition of small amounts of normal plasma restored the prothrombin-time to normal. A somewhat similar new factor is described by Seegers as Ac-globulin. The matter is still not settled satisfactorily.

---

* MELLANBY and PRATT. *Proc. R. S.*, 1940, **128B**, 201.

† WARE *et al.* *Archiv. Biochem.*, 1947, **13**, 231.

‡ WARNER, BRINKHOUS and SMITH. *Amer. J. Physiol.*, 1936, **114**, 667.

The usual source of prothrombin is oxalated plasma. By an elaborate combination of isoelectric precipitation, adsorption on magnesium hydroxide, elution, dialysis, salt fractionation, etc. a pure preparation of prothrombin has been obtained † as a snow-white water-soluble glucoprotein. Pure prothrombin is not altered by calcium ions, but if thromboplastin is also present the prothrombin is converted into thrombin, the Ca ions acting as a catalyst.

For maximal speed of conversion a further factor of a catalytic character is requisite. This is one of the serum globulins, called the Ac-globulin,‡ and it is formed from a plasma Ac-globulin by the action on it of thrombin :—

(1) Prothrombin and Thromboplastin $\xrightarrow{(\text{Ca}^{\cdot\cdot})}$ Thrombin.

(2) Plasma Ac-globulin $\xrightarrow{(\text{Thrombin})}$ Serum Ac-globulin.

(3) Prothrombin and Thromboplastin $\xrightarrow[(\text{Serum Ac-globulin})]{(\text{Ca}^{\cdot\cdot})}$ Thrombin,

The blood of new-born infants is poor in prothrombin, as is also the blood in obstructive jaundice and in vitamin K lack. Administration of vitamin K, with bile salts, by mouth, restores clotting time of the blood to normal. It is therefore believed that the presence of vitamin K is essential to the formation of prothrombin, which is believed to be manufactured in the liver. When dicoumarin is administered, it competes with the vitamin K, and the plasma prothrombin falls, because the circulating prothrombin only has a life of two or three days, and new formation of prothrombin is suspended.

THROMBOPLASTIN. If oxalate plasma be cooled to 0° C. and centrifuged at high speed, a scanty granular precipitate is produced. This consists largely of *débris* from platelets. It is then found that the filtered plasma no longer clots on recalcification, although it still contains fibrinogen and prothrombin. But if the precipitate be treated with calcium chloride and the mixture added to the filtered plasma, the latter clots (although the precipitate freed from adherent contamination will not clot pure fibrinogen). We must conclude that the precipitate contains no thrombin, but will generate thrombin on treatment with prothrombin and calcium salts. It contains the activator, *thromboplastin* (sometimes called "*thrombokinase*").

Oxalate plasma, separated from the precipitate of platelets, can also be made to coagulate by the addition of calcium salts together with extracts of almost any animal tissues, and these therefore also contain thromboplastin. For use as thromboplastin it is usual to employ an extract made from fresh brain or lung §; placenta is also rich in it.‖ These extracts, even on admixture with calcium, are without effect on pure solutions of fibrinogen, since they contain no prothrombin. But if the thromboplastin be treated with prothrombin solution and calcium salts, the mixture when added to fibrinogen causes speedy clotting.

Citrated or oxalated blood can also be caused to clot by the addition of trypsin, and by certain proteolytic snake venoms, which seem to act as thromboplastin. It may indeed be that the thromboplastin-Ca complex is of the nature of a proteolytic enzyme.* The Russell viper venom is a convenient source of thromboplastin for estimations of prothrombin.†

---

* OWREN. *Lancet*, 1947, *i*, 446. *Biochem. J.*, 1948, **43**, 136.

† SEEGERS *et al.* *Archiv. Biochem.*, 1945, **6**, 85 : *J. Biol. Chem.*, 1948, **174**, 565.

‡ WARE, MURPHY and SEEGERS. *Science*, 1947, **106**, 618; WARE and SEEGERS. *Amer. J. Physiol.*, 1948, **152**, 567.

§ SEEGERS *et al.* *J. Biol. Chem.*, 1938, **126**, 91, give further details and also describe preparation of a highly active thrombin.

‖ SCHNEIDER. *Amer. J. Physiol.*, 1947, **149**, 123.

It was shown by Howell that the phosphatid *cephalin* is an essential constituent of thromboplastin, in which it probably occurs in association with a protein. Similar substances are present in nearly all tissues, and are set free when the cells are injured. Separation of the blood platelets by cooling and filtration abolishes the spontaneous coagulability of any form of plasma by removing the source of thromboplastin.

Later work has suggested that there are factors related to thromboplastin in the globulin fraction of normal plasma, and that it may be such factors, "plasma-thromboplastins", associated with platelets that normally initiate clotting. Probably both are needed.

Sequence of Changes. Thus far practically all workers on the subject are agreed. Coagulation of the blood is due finally to the interaction of thrombin and fibrinogen. Fibrinogen and prothrombin are present as such in the circulating plasma. Thrombin is not free in the circulating blood, but is produced or liberated after the blood has been shed. For the production of thrombin from prothrombin the presence of thromboplastin and of calcium ions is necessary.

It is therefore concluded that when blood leaves the vessels there is a disintegration of the blood platelets, with the liberation of thromboplastin. This acts upon prothrombin in the presence of calcium ions and produces thrombin. By the intermediation of the thrombin, the fibrinogen present as a colloidal solution in the plasma is converted into a quasi-crystalline gel of fibrin. The changes occurring in shed blood and resulting in the production of a clot are therefore mainly concerned with the production of thrombin.

This view of the essential character of coagulation is borne out by observations on other forms of plasma, especially of that of birds' blood. This, when obtained with scrupulous cleanliness, so as to avoid any contamination with dust or with the tissues, remains permanently uncoagulable. In the plasma got by centrifuging the blood, no blood platelets are to be seen, and no precipitate is produced by exposure to a temperature of 0° C. We may say, therefore, that blood platelets with their contained thromboplastin have been removed from the plasma, and with them the property of spontaneous coagulability. Since it contains prothrombin as well as soluble calcium salts, it is only necessary to add thromboplastin, *i.e.* a tissue extract, in order to cause immediate clotting.

In every case the initiation of the act of clotting would seem to depend on the addition of thromboplastin to the plasma. In mammalian blood there is no appreciable disintegration of the red or white corpuscles when the blood leaves the vessels, and the first act in coagulation is probably the disintegration of the blood platelets, with the setting free of thromboplastin.

Clotting in Arrest of Hæmorrhage. The part played by the platelets in coagulation is of great importance in maintaining the integrity of the vascular system. If an injury be made to the wall of a blood-vessel which is kept under observation by the microscope, it will be seen that the blood, as it flows past the injured spot, deposits blood platelets on the site of the puncture. These aggregate to form a white thrombus as a plug closely adherent to the wall of the vessel, and effectively preventing any escape of its contents. The white thrombus is almost entirely composed of agglutinated platelets, held together by fibrin. The stickiness of platelets is increased after hæmorrhage or surgical trauma. Another important factor in hæmostasis is *clot retraction* or shrinkage : this also depends on the presence and numbers of platelets. Retraction will draw together and seal off the walls of cut vessels. Later, by the growth of the adjacent endothelial cells, the thrombus is organised and covered with a layer of endothelium continuous with the rest of the vessel. The same process occurs when any part of the lining membrane of a large vessel is injured or when the blood is brought into contact with foreign matter, *e.g.* if a glass or cellophane tube is inserted in the course of a blood-vessel ; if a fine scratch is made inside the tube, platelets congregate speedily in the scratch, and then spread beyond it. Thus, destruction of a patch of endothelium in a vein leads to the deposition of blood platelets over the patch, and the formation of a " white thrombus " adherent to the wall and tending to spread down stream.* Under healthy conditions the thrombus serves simply to cover the bare area in the wall of the vein, and is grown over later by endothelium, so restoring the integrity of the vessel wall. The formation of the white thrombi can be prevented by adequate slow injections of heparin (Best, *et al.*).

---

* EAGLE. *J. Gen. Phys.*, 1937, **20,** 543.

† FULLERTON. *Lancet,* 1940, *ii,* 195.

An important factor in the arrest of hæmorrhage is the increased coagulability of the blood, which clots more rapidly after a considerable fraction of the circulating blood has been withdrawn. Several causes contribute to this, but the most important is probably an increase in the fibrinogen content of the blood. We have seen that fibrinogen is manufactured in the liver, and it would appear that the rate at which the liver adds fibrinogen to the blood is accelerated after a large hæmorrhage. One factor in producing this effect is stated to be the presence of an increased amount of adrenaline in the circulating blood ; as soon as the arterial pressure begins to fall owing to loss of blood the suprarenals discharge adrenaline into the circulation, and this is said to accelerate the coagulation rate of the blood by an action on the liver.† If the liver is excluded from circulation the injection of adrenaline is without influence on the coagulability of the blood, and it also has no effect on it *in vitro.*

Why does Blood not Clot in the Vessels ? No theory of coagulation can be satisfactory which does not account for this. One factor, at any rate, must be the nature of the surfaces with which the blood comes in contact. Blood can for a time be prevented from clotting if it be kept from contact with any foreign substance which is wetted by it, as, for instance, when it is received into paraffined vessels. On the other hand, free contact with such substances hastens coagulation. One must therefore conclude that mere contact with a foreign body has a direct destructive action on the blood platelets. In birds, the platelets are more stable than in mammals, so that coagulation is held in abeyance if contamination with thromboplastin from the juices of damaged tissues is avoided. But birds' blood can be made to clot without the addition of tissue juice if, by violent whipping, or addition of water or chloroform, we destroy the formed elements, so leading to the liberation of their contained thromboplastin.

The modern tendency has, however, been to attribute the fluidity of blood in the vessels to the operation of an anti-coagulation factor, such as the presence of heparin, in the blood.

The blood contains an antithrombin, present in the albumin fraction. Hence thrombin, when incubated with serum or plasma, is partly inactivated. It is not clear what the function of this antithrombin may be ; one suggestion is that it neutralises the small amounts of thrombin formed in the circulating blood.

HEPARIN, when tried on pure constituents (prothrombin, fibrinogen, thrombin), *in vitro*, has no anticoagulant action : it neither inhibits the conversion of prothrombin to thrombin, nor neutralises thrombin or thromboplastin. Yet in blood it acts as a powerful anticoagulant, prevents the conversion of prothrombin into thrombin, and also acts as an antithrombin.‡ For the exercise of these properties, a co-factor, which appears to be a serum albumin fraction, is essential.

---

* BEST, COWAN and MACLEAN. *J. Physiol.*, 1938, **92,** 20.

† GRAY and LUNT. *Amer. J. Physiol.*, 1914, **34,** 232 ; CANNON and MENDENHALL, *ibid.*, p. 243.

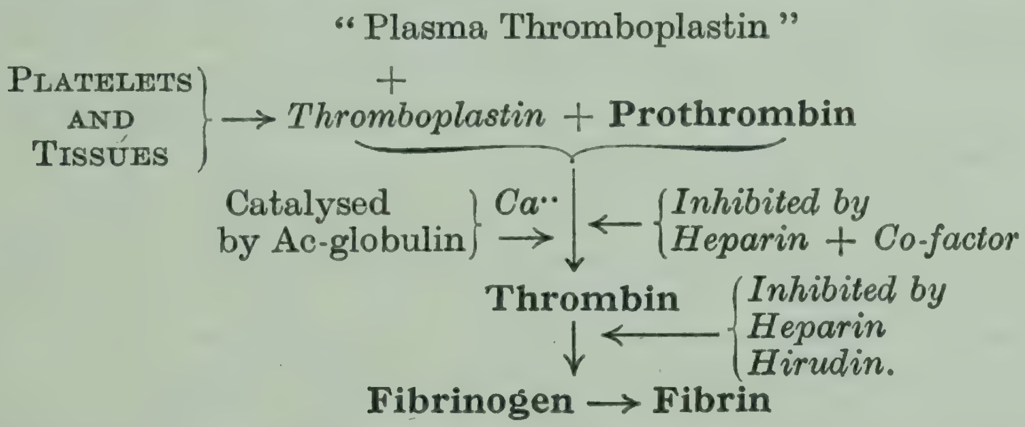
‡ MELLANBY. *Proc. Roy. Soc.*, 1934, **116B,** 1.

By reason of its sulphuric and carboxyl groups, heparin is a fairly strong acid, and combines with proteins to form salts, which are less dissociable the more basic the protein. Some such loose compound may be the real antagonist to the clotting changes in blood: it may even be that the normal blood antithrombin, which, like the co-factor, is an albumin, is the only inhibitor concerned, and that heparin catalytically accelerates its action, though not itself an anticoagulant.

It is doubtful whether heparin is present in normal blood, but there is good evidence that it is formed in the Ehrlich's mast cells, which are found abundantly along the course of many blood vessels. These cells contain granules which stain purple with toluidine blue, a dye that precipitates heparin, and the intima of blood vessels yields, on extraction, quantities of heparin, which are proportional to the abundance of the Ehrlich cells. Under certain conditions the heparin is rapidly poured out into the blood stream, *e.g.* when peptone or small amounts of thrombin or thromboplastin are injected intravenously. In these circumstances not only does no coagulation take place, but the blood obtained after the injection will not clot. Agglutination of platelets is also inhibited by heparin.

Heparin, if injected, is rapidly excreted in the urine, and destroyed in the tissues. Incoagulability *in vivo* can therefore only be maintained by continuous, or frequent, administration.

The following scheme summarises the occurrences in clotting:—



INTRAVASCULAR CLOTTING. Intravascular clotting can be produced by the intravenous injection of small amounts of thrombin: thus 1 mg. of the solid precipitated from the thrombin solution made by Mellanby's method, when quickly injected into the cat's circulation, causes extensive clotting of blood in the vessels and immediate death. Intravascular clotting is still more easily caused by the injection of thromboplastin, *e.g.* of normal saline extracts of tissues rich in cells, such as the thymus, lymph glands or testis.*

A curious result follows from injecting small amounts of these tissue extracts, or larger amounts very slowly. A minute quantity produces, not intravascular clotting, but a *delay* in the coagulation time. Repeated injections of small doses may absolutely annul the coagulability of the blood. The same double effect may be observed even with a larger dose. In rabbits, and in dogs after a full meal, the intravascular coagulation which occurs is complete, extending through the whole vascular system. If the injection be made into a fasting dog the thrombosis produced is limited to the portal vein. These two effects, namely the coagulation and the loss of coagulability, have been called the *positive* and *negative phases* respectively.

At least two different explanations have been given for the occurrence of the negative phase. Since it has not been observed *in vitro* it must be due to a reaction on the part of the living cells and probably to a secretion of heparin into the blood from the mast cells of the blood vessel walls, particularly in such organs as liver or lung. Another explanation has been suggested by Mellanby,† namely, that when small or slow injections of thromboplastin are given, intravascular coagulation actually does occur; the fibrin formed is, however, taken up by tissue cells or slowly deposited on the walls of the blood vessels, so that these are not obliterated; the blood then is really defibrinated blood, and so, of course, will not coagulate. Further evidence in support of this opinion has been given by the observation that slow injections of thrombin give exactly the same results * ; here also the slow injection causes slow continuous coagulation, but the fibrin so formed is safely deposited in places where it can do no harm. There is no reason why both these explanations may not be true in part.

---

* WOOLDRIDGE. "Chemistry of the Blood," 1893. (Kegan Paul, London.)

† MELLANBY. *J. Physiol.*, 1909, **38**, 474.

Fate of the Thrombin. The substances which interact for the production of thrombin and thrombin itself, are not entirely used up in the process of clotting. Blood serum contains traces of thrombokinase (which can be precipitated by the addition of dilute acetic acid) and prothrombin, as well as a fairly strong solution of thrombin. Thrombin, however, rapidly disappears from serum, which in two or three days may be almost free from it. It has been suggested that the thrombin undergoes a modification into an inactive form, called *metathrombin*. According to Rettger the disappearance of thrombin from serum is due to its combination with some of the proteins of the serum. This combination, like that of thrombin with fibrinogen to form fibrin, is unstable and can be broken up by the action of alkalies, acids, or by putrefaction.

We are now in a position to explain the behaviour of the various kinds of plasma.

Cooled Plasma contains the thromboplastin in the blood platelets. These can be separated by centrifuging at a low temperature, or by filtration. The remaining plasma contains only prothrombin, calcium salts and fibrinogen, and can be made to clot by the addition of tissue extracts or of thrombin, but not by warming.

In Sodium Sulphate Plasma the protein factors, or some of them, are precipitated. It is therefore sufficient merely to dilute the plasma in order to produce clotting.

Magnesium Sulphate Plasma if centrifuged while still warm will clot on simple dilution. If the blood be left for twenty-four hours before centrifuging, the plasma will not clot on dilution nor on addition of tissue extracts, because magnesium sulphate slowly precipitates the prothrombin and thrombin.

Sodium Fluoride Plasma might be expected to act like oxalate plasma, but fluoride also causes fixation of the blood platelets. If it be thoroughly centrifuged so that the plasma is obtained free from these constituents, it will no longer clot with calcium salts nor even with calcium salts *plus* tissue extracts, but will clot readily on addition of thrombin. Although it still contains a certain amount of prothrombin, this is adsorbed on the precipitate of calcium fluoride which is produced by the addition of calcium salts, so that the thromboplastin has nothing on which to exercise its effect. Sodium fluoride plasma is therefore useful as a test for the presence of thrombin. If water be added to the sodium fluoride blood, so as to destroy some of the formed elements and liberate their constituents into the plasma, it is possible to produce clotting by the simple addition of calcium salts.

Hirudin Plasma. Hirudin apparently combines with and neutralises thrombin. Hirudin plasma can therefore be made to clot by the addition of thrombin in sufficiently large quantities to over-neutralise all the hirudin.

Heparin Plasma remains fluid because the conversion of prothrombin to thrombin is inhibited. It clots on dilution, or on treatment with carbon dioxide. It also clots on addition of toluidine blue or protamine ; these basic substances withdraw the acidic heparin from its co-factor and precipitate it.

Dicoumarin Plasma is deficient in prothrombin, but not in the labile factor (Prothrombin A or factor 5).

Peptone Plasma. If blood be received into peptone blood, obtained by the injection of large doses of peptone into the veins of another animal, the mixture does not clot. This is because the peptone blood contains an excess of heparin, which can in fact be extracted from it.† It has been shown by Delezenne and others that the seat of its formation is in the liver. In anaphylactic shock also the blood is incoagulable because of the presence in it of an excess of heparin.‡

The Transudations. The earlier work on coagulation was largely carried out on the fluids obtained from the pericardial or pleural cavities, or on hydrocele fluid from the *tunica vaginalis*. These as a rule can be kept indefinitely without clotting, but will clot readily on addition of a few drops of blood or of thrombin. They will not clot on the addition of tissue extracts. Their behaviour is readily explained by the assumption that they contain fibrinogen, but are free from thromboplastin and prothrombin. In order to produce coagulation it is therefore necessary to add both these factors, as happens when we add blood, or else to add thrombin.

THEORIES OF COAGULATION. The incompleteness of our knowledge of what really happens in clotting is reflected in the great variety of theories which have been advanced to explain it. Any theory should have regard to the following facts : (1) Plasma contains all the substances necessary for clotting. (2) Blood does not clot in normal

* Mellanby. *Proc. Roy. Soc.*, 1933, **113B**, 93.

† Wilander. *Skand. Arch. Physiol.*, 1938, **81**, Suppl. XV.

‡ Jaques and Waters. *J. Physiol.*, 1941, **99**, 454.

blood vessels. (3) Coagulation is closely related to the properties of the colloids concerned in it. (4) Essential factors in clotting are primarily fibrinogen and thrombin; secondary factors which govern the formation of thrombin are prothrombin, calcium salts and thromboplastin. (5) In hæmophilia the blood is only slowly coagulable or even incoagulable. Thus far, though with different nomenclatures, most workers are now agreed.

Morawitz called the thromboplastic factor thrombokinase and considered that this substance, liberated by platelets or damaged tissues in presence of calcium salts, converted a precursor, prothrombin, in the plasma, into thrombin.

According to Nolf the essential factors are three proteins : fibrinogen, thrombogen and thrombozym. The two former are produced in the liver, while the thrombozym is formed from the leucocytes. The clotting depends on a mutual interaction and precipitation of colloids with, as a result, either fibrin or thrombin. Thrombin differs from fibrin merely in containing less fibrinogen. For this reaction to take place, the presence of calcium is necessary as well as certain thromboplastic substances which act as centres of precipitation. Thrombin, according to him, is merely an unsaturated compound which is capable of taking up or uniting with more fibrinogen to form fibrin. He compares these actions to the actions of digestive enzymes on proteins ; fibrinogen is first precipitated as fibrin by union with thrombozym and thrombogen. This fibrin is then hydrolysed and dissolved by the further action of the thrombozym, which he regards as essentially proteolytic in character. That the blood does not coagulate within the vessels he explains by assuming that the cells of the blood and tissues are covered normally with an ultra-microscopic layer of fibrin which has no thromboplastic effect upon the plasma.

Howell's theory assumes that the thromboplastic material (cephalin) liberated by platelets or by other damaged tissue cells neutralises heparin, which normally maintains the fluidity of the blood *in vivo*, so that calcium salts can then convert prothrombin to thrombin.

It seems certain that complex physico-chemical processes connected with electric charges on colloids or high-molecular substances play an important part, since many anti-coagulants are strongly acidic or basic substances of large molecular weight. For instance, the lignine sulphonic acids (Liquoid Roche), Germanin and heparin are acidic ; Janus green is basic.

HÆMORRHAGIC STATES. In some pathological states the bleeding from a wound is abnormally prolonged. This may be due to one or more of many causes. One of these conditions is called *hæmophilia.*

In this condition the blood clots, if at all, with extreme slowness ; hence subjects suffering from the complaint may die of hæmorrhage as a result of trivial injuries. Yet their blood undoubtedly contains prothrombin and fibrinogen, so that if thromboplastin solution is added it clots normally. The most probable explanation is that the platelets of hæmophilic blood are less readily made to yield sufficient thromboplastin than those of normal blood, either because of their small number or their relative stability.

Platelet-free plasma, normal or hæmophilic, does not coagulate on contact with glass ; adding normal platelets to hæmophilic blood does not induce clotting either ; but if *normal* plasma is also added, together with *either* normal or hæmophilic platelets, the hæmophilic plasma then clots in glass.* It appears, therefore, that the plasma thromboplastin is lacking, and hence transfusions of normal blood greatly shorten the clotting time in hæmophilic cases. Local application of certain snake venoms to the wound accelerates clotting. The disorder is of further interest because it is hereditary and is transmitted through the females, who are not affected by it, to the male descendants through many generations.

In other cases there is a lack of prothrombin, as in jaundice, or some days after administration of " dicoumarin " or of salicylates, and in certain nutritional disorders due to lack of vitamin K ; or there may be lack of fibrinogen, as in gross liver damage or in the hereditary pseudohæmophilia. Very rarely, if ever, is excessive bleeding due to calcium deficit, because for this to occur the $Ca^{\cdot\cdot}$ would have to be reduced to a level incompatible with life. A primary lack of some clotting factor is, however, not the only cause of excessive bleeding. For instance, in anaphylactic shock the blood is incoagulable because of the presence of an excess of heparin.

Often the hæmorrhagic state is associated with the occurrence of purple or livid spots on the skin—a symptom called purpura—which is due to fragility of the capillary walls. This may be associated with reduced blood coagulability, as in the purpura due to lack of vitamin K, but often the coagulation time is normal and there is no detectable lack of any coagulation factor in the blood. The most outstanding of such examples is scurvy, due to lack of vitamin C ; another may be due to lack of another vitamin, called vitamin P (flavone) so raising the permeability of the blood capillaries. Again, some states of purpura (idiopathic thrombocytopenic purpura) are distinguished by a reduction in the number of blood platelets. Whether this is due to sluggish production or accelerated destruction of platelets is unknown, and the condition is often relieved by removal of the spleen. Platelet deficiency is also present in the purpuras of aplastic anæmia and lymphatic leukæmia. Some believe that the primary fault is in the capillary endothelium, which is excessively vulnerable;

---

* BRINKHOUS. *Proc. Soc. Exp. Biol. Med.*, 1947, **66**, 117.

hence, many platelets are used up to repair the breaches, with consequent lowering of the circulating platelets.

For Reference

BEST (1940). Heparin and Thrombosis. (*Harvey Lects.*, 1940–41; p. 66.)

FERGUSON (1936). The Blood Calcium and the Calcium Factor in Blood Coagulation. *Physiol. Rev.*, **16,** 640.

HOWELL (1935). Theories of Blood Coagulation. *Physiol. Rev.*, **15,** 435.

JORPES (1939). Heparin. *Oxf. Med. Publ.*

MACFARLANE (1948). Normal and Abnormal Blood Coagulation. *J. Clin. Path.*, **1,** 113.

QUICK (1942). The Hemorrhagic Diseases. (Springfield. Charles C. Thomas.)

QUICK (1944). The Anticoagulants effective *in vitro* with special reference to Heparin and Dicumarol. *Physiol. Rev.*, **24,** 297.

SILBERBERG (1938). Causes and Mechanism of Thrombosis. *Physiol. Rev.*, **18,** 197.

WÖHLISCH (1940). Fortschritte in der Physiologie der Blutgerinnung. *Ergebn. d. Phys.*, **43,** 174.

THE CIRCULATION

CHAPTER XXX

GENERAL FEATURES OF THE CIRCULATION

In the tissues, the blood is passing continuously through a close meshwork of capillaries, minute vessels with walls consisting of a single layer of delicate endothelial cells, which permit of a free interchange of material by diffusion between the blood within and the tissue fluid outside the vessel. The movement of the blood is maintained by the heart, which acts as a pump, transferring the blood it receives at the low pressure of the veins, to the arteries at a high pressure, the blood being taken from the heart to the tissues by the arteries, and carried back from the tissues to the heart by the veins.

In all the vertebrates the vascular system is closed, *i.e.* communicates at no point with the tissue spaces or cœlomic cavity. It is found in its simplest form in fishes (Fig. 338, A), where the heart consists of one auricle and one ventricle. The blood is received from the great veins into the auricle. The walls of both auricle and ventricle contract rhythmically. By the contraction of the auricle the blood is forced into the ventricle and this, when it contracts, sends the blood on into the bulbus arteriosus. From the bulbus the blood passes through the branchial arteries into the gills, where it takes up oxygen from the surrounding water, and then flows on into the aorta, by which it is distributed to the various organs of the body. From the capillaries of these organs the blood is collected by the veins and is carried once more back to the auricle. The fish heart is thus entirely on the venous side of the vascular system.

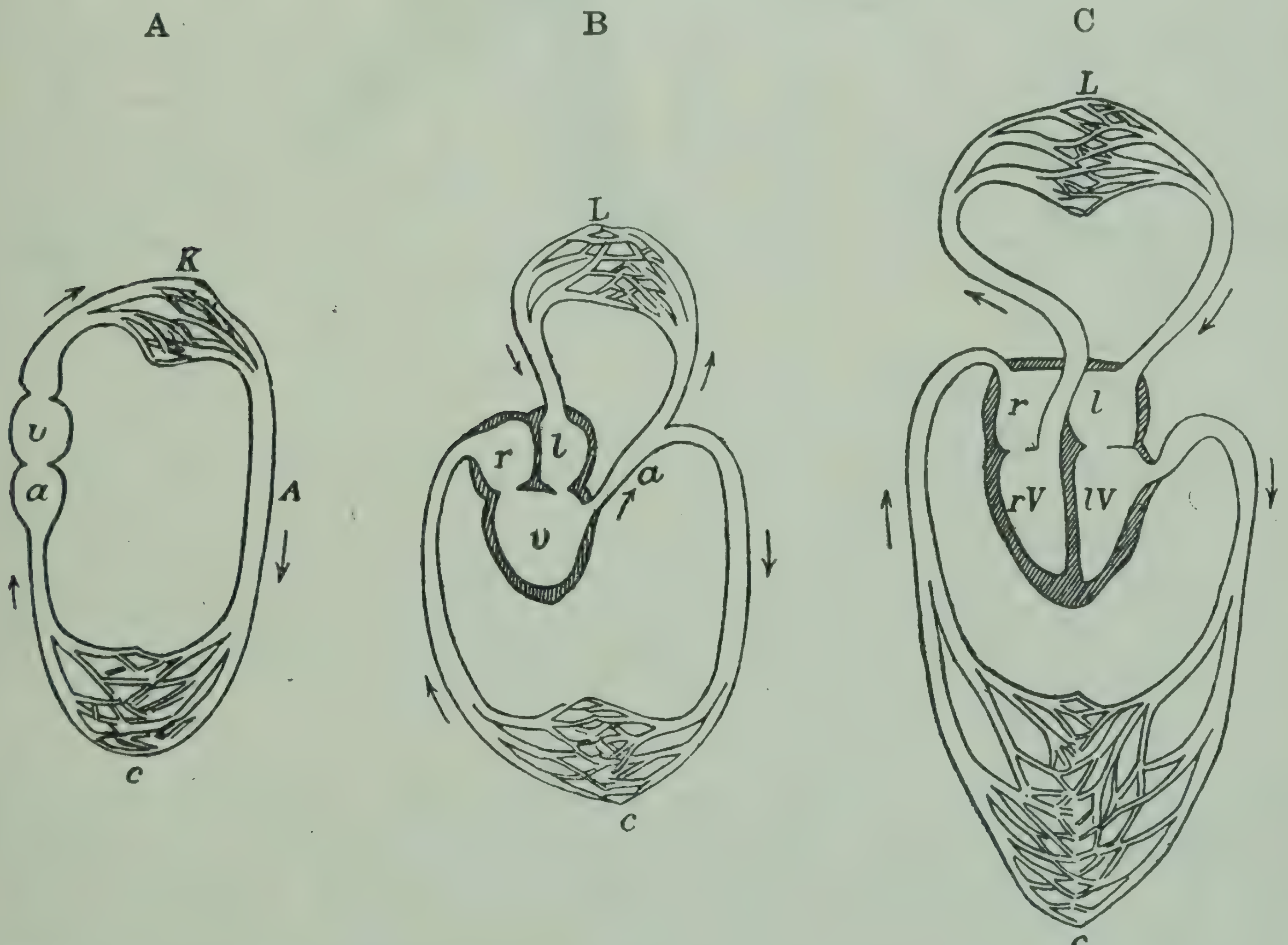


FIG. 338. Diagram of Circulatory System in A, Fish ; B, Amphibian (Frog) ; C, Mammal.

v. Ventricle.
a. Auricle.
K. Gill capillaries.
A. Aorta.
c. Systemic capillaries.
L. Lung capillaries.
r, *l*. Right and left auricles.
rV, *lV*. Right and left ventricles.

In amphibia, such as the frog (Fig. 338, B), the heart consists of two auricles and one ventricle. The right auricle receives venous blood from the body by means of the venæ cavæ and forces it by its contraction into the ventricle. From the ventricle the blood passes into the aorta, whence it is carried partly by the pulmonary artery to the lungs, partly by arteries to the different organs of the body. The blood, which has passed through the lungs and been oxygenated, flows through the pulmonary veins to the left auricle, whence it passes into the ventricle and mixes with the venous blood which is arriving from the right auricle. The pulmonary circulation is thus merely a branch of the general or systemic circulation. The bulbus aortæ in the frog is divided into two parts by means of a spiral valve, by which a partial separation of the blood coming from the right and left auricles is effected, and the venous blood from the right auricle directed especially into the pulmonary artery.

In birds and mammals the heart has become entirely divided into two halves, right and left, which have no communication with one another except by way of the blood vessels and capillaries. The right auricle receives the venous blood from all parts of the body and sends it on to the right ventricle, whence it is forced into the lungs along the pulmonary artery. In the lungs it takes up oxygen and so becomes " arterial " and is returned by the pulmonary veins to the left auricle and on to the left ventricle. The rhythmic contractions of the left ventricle then force the blood into the aorta, whence by the branching arteries it is carried to all parts of the body (Fig. 338, c).

The whole vascular system is distensible and elastic, so that its capacity will increase with the pressure of the blood contained in it. Since the driving force is furnished by the heart, the pressure which causes the flow of blood through the system must decline as we pass from the arterial to the venous side (Fig. 339). In the arteries the pressure is high, *e.g.* about 120 mm. Hg in the large arteries of man ; it is somewhat lower in the smallest arteries. The chief function of the large arteries is to serve as elastic conduits, whereas the small arteries and arterioles leading from the arteries to the capillaries have in addition the function of regulating the amount of blood flowing through the capillary area of the organs which they supply. In the capillaries the pressure is much lower, say, 30 to 50 mm. Hg. The veins have the function of conducting blood at a low pressure (not usually greater than 1 mm. Hg) from capillaries to heart and of holding any excess of blood which is not immediately taken up by the heart.

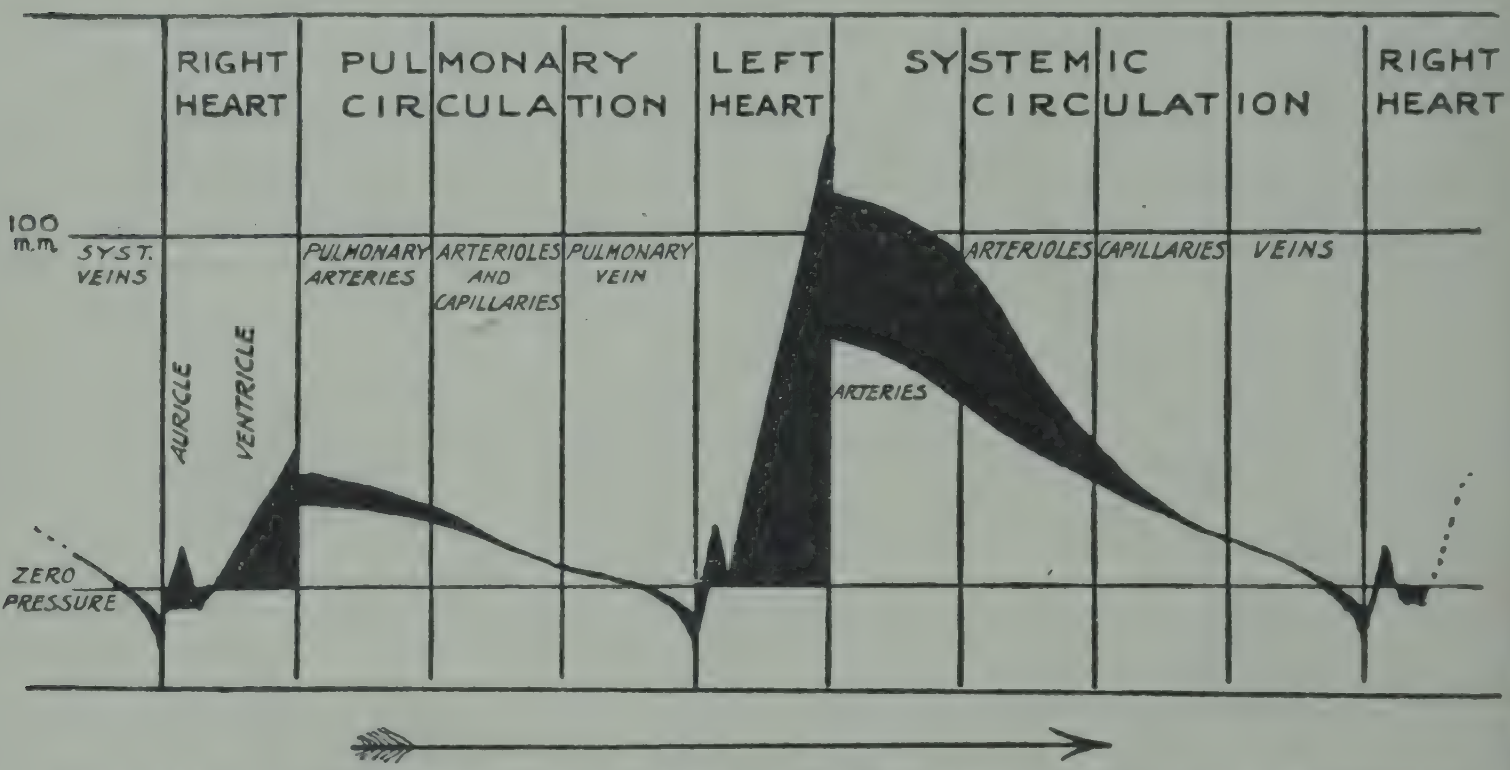


FIG. 339. Diagram to illustrate the pressures in different parts of the vascular circuit. The circulation proceeds in the direction of the arrow. Lines of " zero " (= atmospheric) pressure and of 100 mm. Hg pressure are shown. The pressures are shown as a graph of varying breadth ; the breadth at any point indicates the approximate magnitude of the pressure fluctuations caused by the heart beats. These fluctuations are at a maximum in the ventricles and at a minimum in the smallest capillaries and veins, in which the pressure is almost constant. (The fluctuation in the pulmonary arteries is probably much greater than is here shown.)

Corresponding to this difference in function we find variations in the structure of the blood vessels according to their situation in the circuit. The arteries are thick-walled and contain an abundance of muscular and elastic elements in their walls. The typical medium-sized artery is described as consisting of three coats (Fig. 340): an *intima* lined by a continuous layer of flattened endothelial cells, which rest on a well-marked lamina of yellow elastic tissue; a *media* composed of unstriated muscular fibres arranged mainly circularly; and an external coat or *adventitia* of fibrous tissue, with a number of longitudinal elastic fibres. Near the heart, in the great vessels such as the aorta and its larger branches, there is a preponderance of elastic tissue as compared to the muscular; and we find in the media alternate layers of muscle fibres and fenestrated elastic membranes.

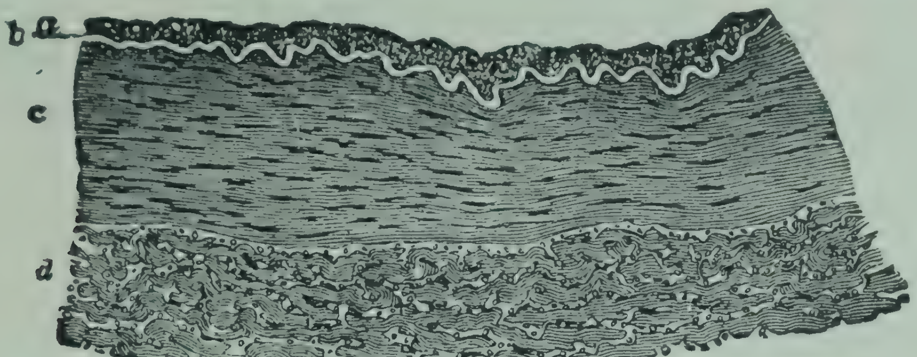


FIG. 340. Transverse Section of Part of the Wall of the Posterior Tibial Artery (× 75). (SCHAFER.)

a. Endothelial and sub-endothelial layers. of intima.
b. Lamina of elastic tissue.
c. Media consisting of muscle fibres.
d. Adventitia.

Arteries are of two types, the "elastic" and the "muscular." The aortic arch and its main branches are of the "elastic" type; they constitute a large elastic reservoir which lightens the work of the heart and converts the intermittent output into a continuous flow. The smaller arteries, such as radial and lingual, are of "muscular" type; they serve to regulate the supply of blood to the various organs as their needs require it. In the smallest arteries, the *arterioles*, the elastic element entirely disappears, so that the wall consists of muscle fibres, chiefly circular, lined by the endothelium. In these vessels a contraction of their walls may result in an entire obliteration of the lumen, so shutting off altogether the supply of blood to the capillaries beyond. The capillaries, though formed merely of a single layer of endothelial cells, also possess the power of contracting and relaxing, and so admitting of a larger or smaller flow of blood to the tissues according to their local needs. In the veins the same three coats can be distinguished as in the typical artery, but the wall of the vessel is much thinner in proportion to the lumen. In the vein, moreover, there is a preponderance of the fibrous tissue elements, the muscular and elastic tissue being but little marked. On this account the vein collapses unless it is distended by some internal pressure.

The histological difference between veins and arteries is of importance for the understanding of the distribution of pressures in the vascular system, since the distensibility of these vessels is conditioned by their structure. In Fig. 341 is represented the distensibility, *i.e.* the increase in capacity, of a vein under gradually increasing internal pressure. It will be seen that the capacity, which is negligible at zero pressure, becomes considerable on raising the pressure to

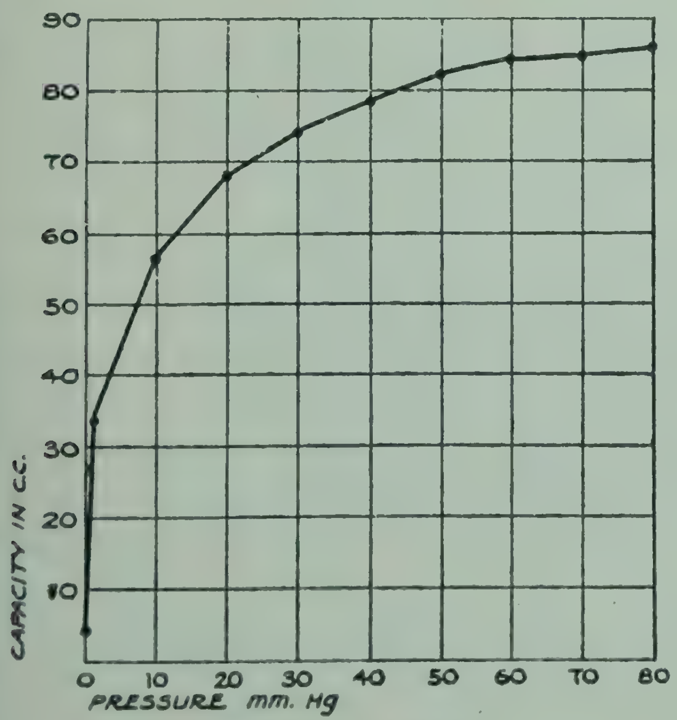


FIG. 341. Vein.

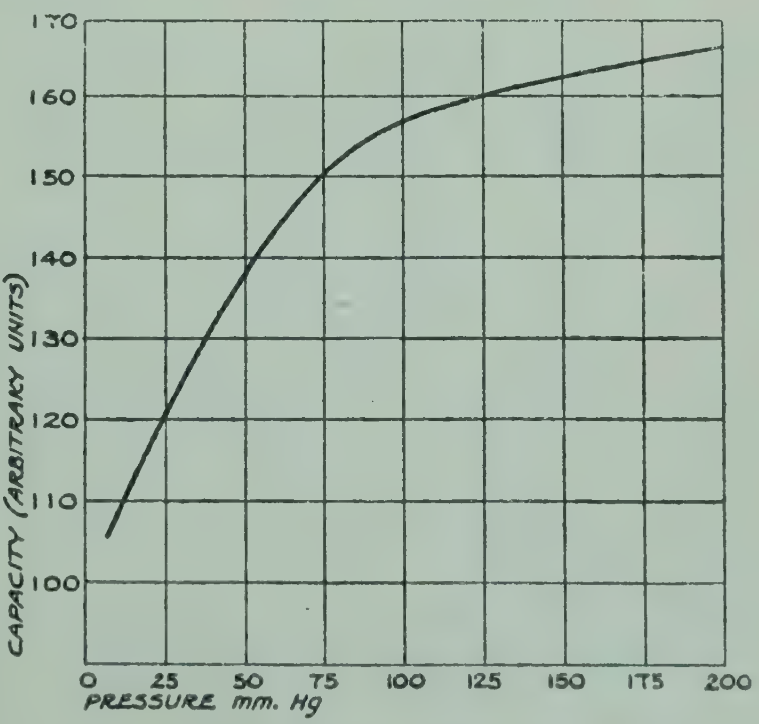


FIG. 342. Artery.

1 mm. Hg. A further rise of pressure to 10 mm. Hg causes a very large increase in volume, but from this point the increments of volume with rising pressure rapidly diminish. On the other hand, the artery, which has a certain capacity at zero pressure, increases its capacity much more slowly with rise of pressure, and the increment of capacity becomes gradually less as the internal pressure rises. In Fig. 342 is shown the distensibility of an artery to sudden alterations of pressure such as occur with each heart beat. It indicates that the arterial wall becomes more rigid with increasing distension. In the thoracic aorta of young subjects the curve is S-shaped, the most rapid distension taking place at about 100 mm. Hg. pressure.* The aorta is thus well suited to expand at each systole.

As the arteries branch, although each branch is smaller than the parent vessel, the total area of the two branches into which the vessel divides is greater. Thus there is a continual increase in the cross area of the bed of the blood stream as we pass from the heart towards the periphery. This increase is especially marked at the junction between the capillaries and the arterioles on one side and the venules on the other, so that the total area of the bed in the region of the capillaries can be taken as about 800 times that of the area of the aorta where the blood leaves the heart.

This figure would apply only to normal conditions, under which but a portion of the capillaries are open and pervious to the blood current. By the opening up of previously closed capillaries, the capacity of the capillaries in active muscle may be increased to many times that of these vessels in a resting muscle. In fact, a general and simultaneous dilatation of all the capillaries of the body may increase the capacity of the stream bed to such an extent that there is an insufficient return of the blood to the heart, and a condition of "shock" is brought about.

Arterio-venous Anastomoses. Although the bulk of the circulating blood has to pass through capillaries in order to travel from the arterial to the venous system, this is not invariably the case, for in certain parts of the body there are direct channels leading from arteries to veins. These are called arterio-venous anastomoses. In them the arterial twig, after a more or less tortuous course, opens directly into a relatively wide, thin-walled vein. The structures are often collected together in groups in certain situations, or may be scattered singly. They occur especially in erectile tissue, such as the penis, in the skin of the balls of the fingers, in the nail bed, and to a remarkable degree in the rabbit's ear skin. The openings of these anastomoses are capable of being closed by the contraction of a thick muscular wall; when this is relaxed the arterial blood is short-circuited directly into the vein.

For Reference

CLARK (1938). Arterio-venous Anastomoses. *Physiol. Rev.*, **18**, 229.

---

* HALLOCK and BENSON. *J. Clin. Investig.*, 1937, **16**, 595; WIGGERS. *Amer. J. Physiol.*, 1938, **123**, 644.

CHAPTER XXXI

PHYSIOLOGY OF THE HEART

CARDIAC MUSCLE

In the ordinary span of a man's life his heart performs about 2,600 million beats, pumping at least some 155 million litres (say about 150,000 tons) of blood from each ventricle: at the lowest estimate the work done amounts to 241 million kg.-m., enough to raise a weight of 10 tons to a height of ten miles. In all this performance scarcely a beat is misplaced, and there is no rest longer than about three-quarters of a second throughout. It must be admitted that this is a remarkable performance for an organ weighing only about 300 grm., even when spread over seventy years.

If the heart be cut out of the body of a cold-blooded animal it will continue to beat with the normal sequence of its different chambers for hours, or even days, provided that it be kept cool and moist. In the case of a warm-blooded animal, the heart is similarly capable under proper conditions of continuing its rhythmic contractions for many hours after excision. The fact that the heart will continue to beat after removal from the body shows that the causation of the beat is to be sought in the walls of the heart itself. The heart wall consists of cardiac muscular tissue; it presents longitudinal and transverse striations and is capable of contracting both "spontaneously" and in response to direct stimulation. The histological unit is a muscle cell, roughly quadrilateral when

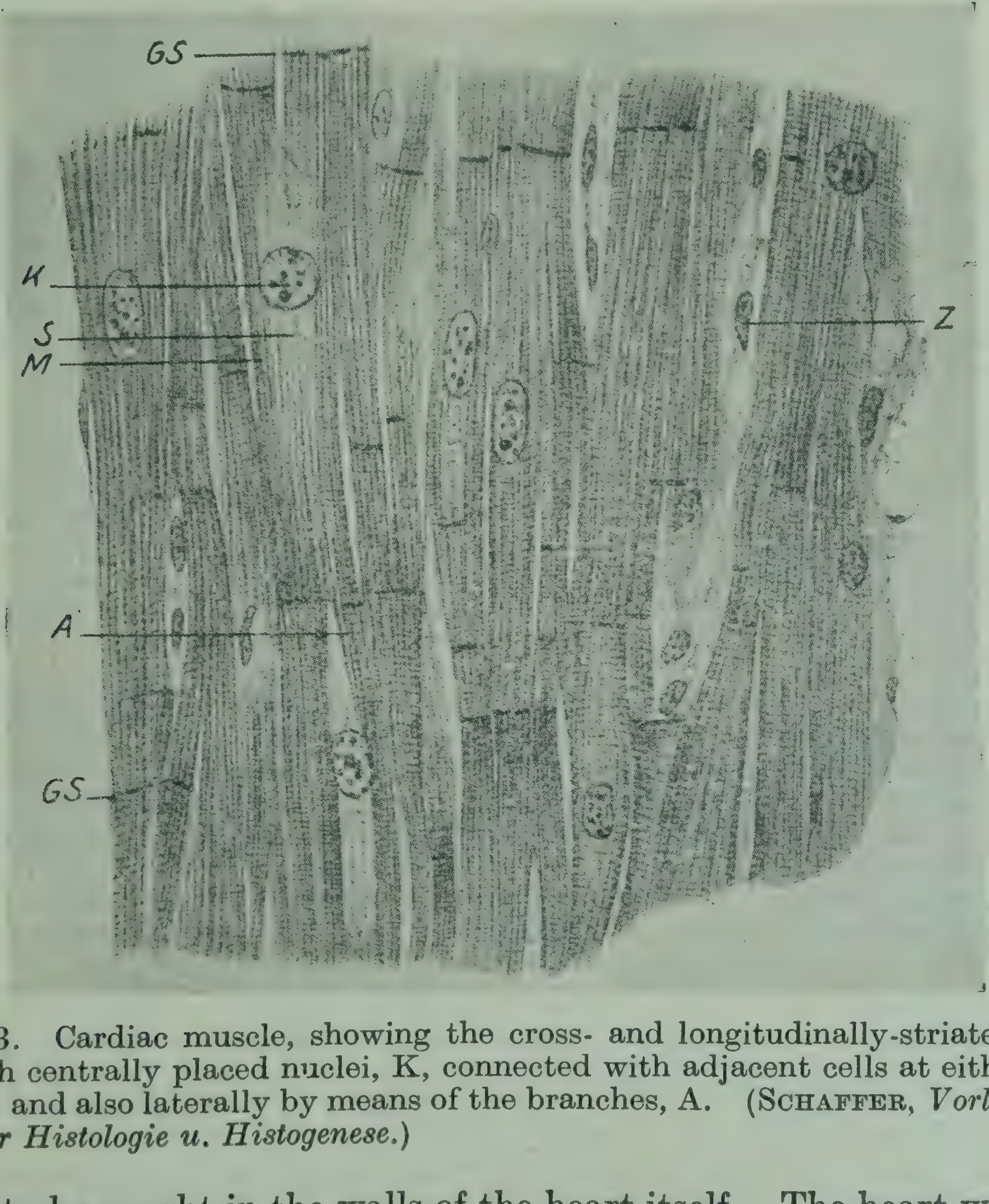


FIG. 343. Cardiac muscle, showing the cross- and longitudinally-striated cells, with centrally placed nuclei, K, connected with adjacent cells at either end, GS, and also laterally by means of the branches, A. (SCHAFFER, *Vorlesungen über Histologie u. Histogenese.*)

viewed from the surface. Each cell has a nucleus near its centre, and is joined to adjacent cells to form fibres. No sarcolemma is present, and the fibres lying side by side are connected at frequent intervals by short bridges (Fig. 343). The contractile units are thus not separate, as in voluntary muscle, but are in protoplasmic and functional continuity, *i.e.* they form a *syncytium*.

ANATOMY OF THE HEART IN THE FROG AND TORTOISE. The hearts of the frog and of the tortoise have figured largely in the researches on the causation of the heart beat.

The frog's heart (Fig. 344) consists of the sinus venosus, which receives the anterior and posterior venæ cavæ, two auricles, one ventricle, and the bulbus arteriosus which opens into the two aortæ. The venous blood from the body flows into the sinus venosus by the three venæ cavæ, and thence into the right auricle, while the left auricle receives the blood from the lungs. The ventricle thus receives mixed arterial and venous blood, the arterial blood being directed by the spiral valve of the bulbus aortæ so as to flow chiefly towards the head.

The muscular fibres of the frog's heart are less highly developed than those of the mammalian heart. They are spindle-shaped, and only dimly cross-striated. The cross-striation becomes more distinctly marked as we proceed from sinus to ventricle, the sinus muscle fibre representing the most primitive condition. There is complete muscular continuity between all the cavities of the heart. The circular ring of muscle at the junction of sinus with auricles and of auricles with ventricle presents only slight traces of cross-striation.

The heart is well supplied with nerve fibres and ganglion cells. The two vagi enter the sinus venosus and branch just under the pericardium. Here they become connected with a collection of nerve cells, known as Remak's ganglion. From the sinus the two vagi, now called septal nerves, pass down in the interauricular septum, one in front and the other behind. Near the auriculo-ventricular groove they enter two collections of ganglion cells, called Bidder's ganglia. From these ganglia non-medullated fibres are distributed to surrounding parts of the auricle and to the whole of the ventricle. In the upper third of the ventricle occur scattered ganglion cells attached to the nerve fibres. These are quite absent in the lower half or two-thirds.

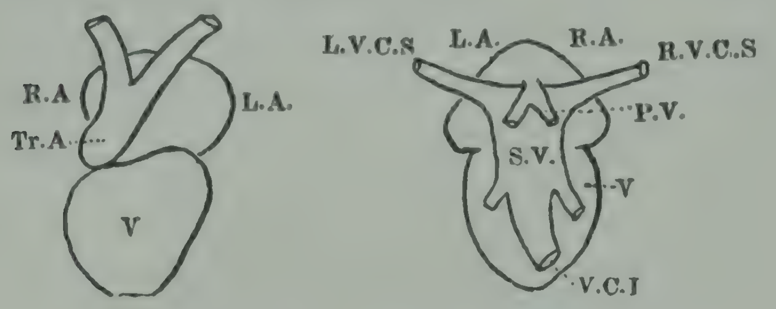


FIG. 344. Diagram of Frog's Heart. (After CYON.)

V, ventricle; R.A., L.A., right and left auricles (atria); S.V., sinus venosus; P.V., pulmonary veins; L.V.C.S. and R.V.C.S., left and right anterior vena cava; V.C.I., vena cava posterior; TR.A., bulbus arteriosus.

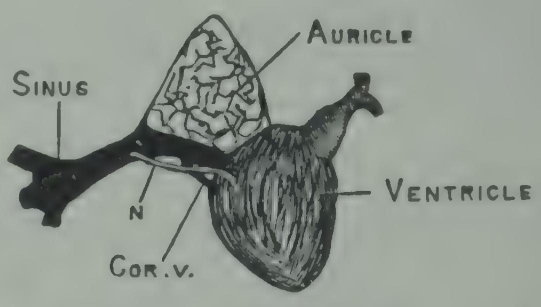


FIG. 345. Tortoise's Heart (after GASKELL) as it appears when Suspended for Registering the Auricular and Ventricular Contractions. N, nerve trunk with fibres connecting Remak's and Bidder's ganglia; COR. V., coronary vein.

In the tortoise (Fig. 345) the two auricles are bound together by a flat band of tissue, which serves also to connect the sinus with the ventricle. The septum between the auricles arises from the central line of this junction wall. The two vagus nerves pass into a large accumulation of ganglion cells in the sinus, and thence along the basal wall to the auriculo-ventricular groove, lying just under the pericardium. In the groove they pass into a collection of ganglion cells, whence fibres are given off to both auricles and ventricle. As they leave the sinus, a branch of the right nerve runs along the coronary vein, which conveys blood from the ventricular wall to the sinus.

The Contraction of the Cold-blooded Heart. THE STANNIUS LIGATURES. The frog's heart beats regularly, the contraction starting in the sinus, then travelling to auricles, ventricle and bulbus. If, however, the auricles be functionally separated from the sinus by a ligature round this junction (first Stannius ligature), the auricles and ventricle stop in an uncontracted condition (diastole), while the sinus goes on beating regularly. After the lapse of a period varying from five minutes to half an hour, the detached part of the heart begins to beat, at first slowly and then more rapidly, but never attaining the rate of the sinus. The auricles beat first and then the ventricle. If now the ventricle be separated from the auricles by a ligature at the auriculo-ventricular groove (second Stannius ligature), the auricles continue to beat, while the ventricle, after a few beats due to the excitation of the ligature, stops beating, though after a considerable time it may begin again to contract very slowly. But the lower two-thirds of the ventricle, if separated from the rest of the heart, never beats again, under normal circumstances. To single stimuli it responds with a single beat, not with a series of beats as the whole heart does. It can be made to beat rhythmically, however, by raising the pressure within its cavity.

There is thus a descending scale of automatic power in the different parts of the frog's heart—from the sinus, where it is highest, to the lower parts of the ventricle, where it is very slight. The normal sequence of events—*i.e.* the subordination of the ventricle to auricles, and auricles to sinus, so that the beat always follows in the order, sinus, auricles, ventricle, bulbus—can be ascribed to the difference between the natural rhythms of these different cavities. It is possible to record the contractions of each of these parts of the heart separately, after having divided them, either functionally by crushing the intervening tissue, or by actual section. Under such conditions it is found that there is a descending scale of rhythm from sinus to bulbus, the contractions of the sinus being most frequent, those of the ventricle and bulbus the least frequent. Thus in the entire heart, it is impossible for the ventricle to function at its own rhythm, since before it is ready to beat again spontaneously, it receives an impulse from the auricles which induces a beat. That the normal sequence of contractions is dependent on the domination by the sinus, is shown by the fact that, by exciting the ventricle by means of induction shocks repeated at a rhythm slightly quicker than that of the sinus, it is possible to induce a reverse rhythm, the order of the beat being now ventricle, auricles, sinus venosus. It is clear, therefore, that the portion of the heart which beats most rapidly dominates the rest, and it is therefore called the *pace-maker*.

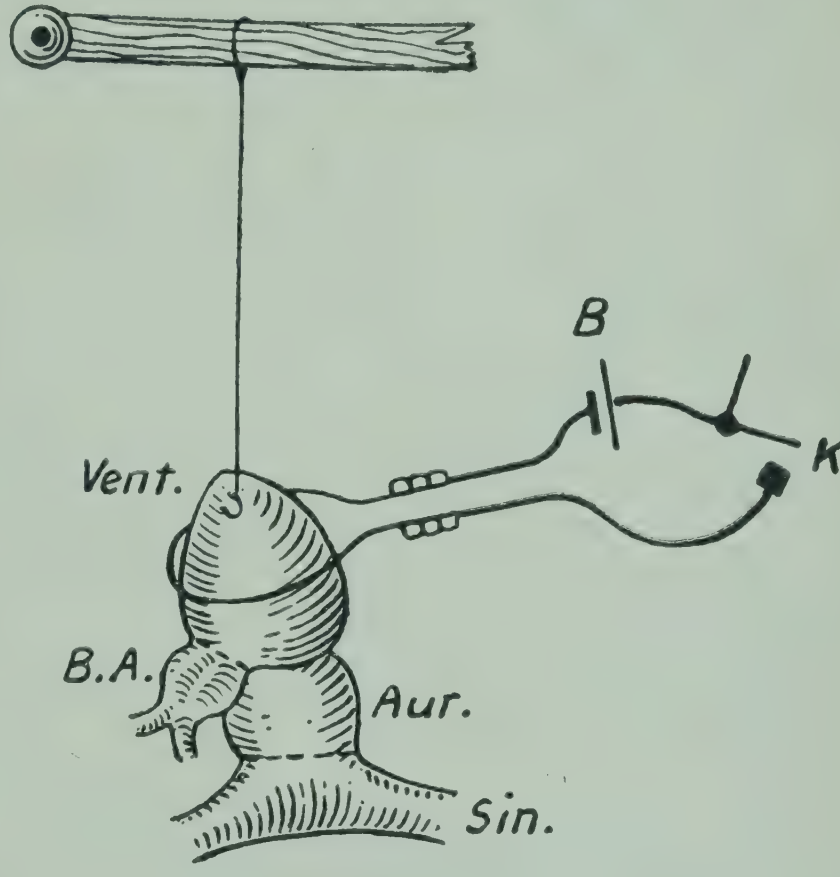


FIG. 346.

That the sinus is the pace-maker in the frog's heart may be shown by a simple experiment. The ventricle is connected with a recording lever, and a platinum loop, which can be heated by passing a current, is put round the heart (Fig. 346). When the ventricle alone is warmed the rhythm is unaltered. On lowering the loop so as to warm the sinus, the rhythm of the whole heart is quickened.

The different rhythmic power of these parts of the heart is apparently connected with the histological characters of the muscle fibres at each part. The lowly differentiated sinus cell has well-marked rhythmic power and a quick rhythm of beat, but is not able to exert much force in its contraction. The more highly differentiated ventricle cell has only a slight rhythmic power, but beats forcibly.

NEUROGENIC THEORY OF HEART BEAT. It was formerly thought that the ganglia (Remak's and Bidder's ganglia) were responsible for the initiation and co-ordination of the beats. This theory has been abandoned since it was shown by Gaskell and by Engelmann that—

(*a*) These ganglia may be excised in the frog and tortoise without altering the contractions or normal sequence of beat.

(*b*) Remak's ganglia are merely distributing ganglia for the vagus nerves to the heart. They may be paralysed by nicotine without disturbing the beat.

(*c*) A portion of cardiac muscle, free from ganglion cells, can contract rhythmically.

(*d*) The heart in the developing chick begins to contract at a date prior to the growth into it of nerve cells.

Finally it has been shown that cardiac muscle cells in tissue culture show rhythmic contractions.

THE PROPAGATION OF THE WAVE OF CONTRACTION. The normal contraction is started in the sinus venosus, propagated to the auricles, thence to the ventricle, and finally to the bulbus aortæ. Between the contractions of each of these cavities there is a slight pause, whereas the contraction spreads so rapidly over each cavity that all parts, say of the auricles or ventricle, appear to contract simultaneously. Cardiac muscle forms a syncytium, and the excitatory wave is propagated through the heart from one muscle cell to another, and not by means of nerve fibres.

That the propagation cannot be due to any nerve trunks running from sinus to ventricle is shown by various experiments of Engelmann and Gaskell.* Thus, if the auricle is slit up by a series of inter-digitating cuts, the contraction wave starting from the sinus travels along the auricular muscle around the end of each section and finally, on arrival at the ventricle, causes a contraction of this cavity. In the heart of the tortoise, the nerve trunks run in the band of tissue joining the sinus to the ventricle ; this band, with all its contained nerves, can be excised without interfering in any way with the normal sequence of contractions.

By the methylene blue method it is possible to demonstrate a close network of non-medullated fibres surrounding all the muscle cells of the heart. It is obvious that some known facts would not exclude the possibility of propagation occurring through such a nerve network, but the properties of the network would have to differ from those of any of the nerve tissues with which we are acquainted ; for example, the velocity with which the excitation is forwarded is much slower than that in nerve.

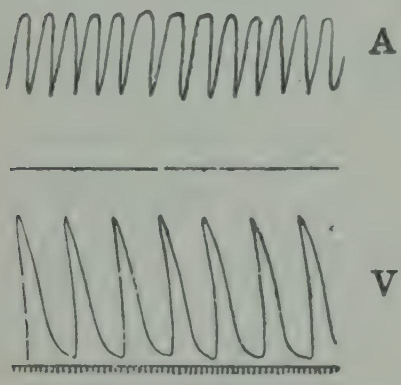


FIG. 347. Contraction of Auricles and Ventricle of Tortoise Heart. The auriculo-ventricular groove has been clamped so as to produce a partial block allowing only every second contraction to pass. (GASKELL.)

The pause between the contractions of auricles and ventricle, according to the more modern opinion, first advanced by Erlanger, is due to a latency in response of the junctional tissue to the action potentials reaching it from the auricle.† A similar retardation of the wave can be produced at any point, in either auricles or ventricle, by compression, or by cutting away the conducting muscular tissue to a sufficiently small amount. If the normal auriculo-ventricular " block " be increased, as for instance by compression of the auriculo-ventricular groove, it may be that only one out of every two or more auricular contractions passes across to the ventricle (Fig. 347), and further increase in the resistance to the propagation of the wave may lead to the block becoming complete.

The progress of the excitatory wave is well seen if a record be taken of the electrical changes resulting in the frog's heart from a single stimulation. A series of electrical changes occurs, which can be interpreted as due to the passage of a wave of excitation from the base to the apex of the heart.

THE PROPERTIES OF CARDIAC MUSCLE

Rhythmicity. The cause of the spontaneous rhythm of the various portions of the heart, and of the higher rhythm of the sinus tissue, is not known. It is evident that some stimulating agent is present in the cardiac cells themselves, and attention has been given to the possibility that this stimulus is a chemical one, as in the " heart hormone " of Haberlandt ‡ or the

---

* GASKELL. *J. Physiol.*, 1883, **4**, 43.

† GIBSON. *Amer. J. Physiol.*, 1942, **138**, 113.

‡ HABERLANDT. " Das Hormon der Herzbewegung." 1927.

" active substance " of Demoor.* The higher rhythmic activity of the sinus is, according to these views, due to the more rapid production of the " hormone " in that region. According to others it is the outcome of a more vigorous metabolism.

Response to Direct Excitation. The curve of contraction of cardiac muscle resembles that of voluntary muscle, except that it occupies a longer time ; the contraction of the mammalian ventricular muscle lasts three-tenths of a second, of the frog's ventricle about half a second, and of the tortoise ventricle about two seconds. The contractile process originates at a stimulated point and travels thence to all other points.

All-or-none Contraction. When a ventricle strip is stimulated with a single induction shock, if it responds at all it will respond with a maximal contraction, no increase being obtainable, however the stimulus may be increased.† There is thus no proportionality in the heart between strength of stimulus and force of contraction. We have seen that the response to a minimal stimulus in a skeletal muscle is smaller than the response to a maximal stimulus, because in the former case only a small proportion of the muscle fibres is active, so that increasing the strength of the stimulus merely increases the number of fibres thrown into contraction, and a maximal contraction of skeletal muscle is one involving all the fibres. In the heart muscle all the muscle fibres are functionally continuous, so that a stimulus, if it excites at all, must excite all the fibres, and every contraction must be analogous to the maximal contraction of a skeletal muscle. The " all-or-none " law is therefore demonstrable without special methods in heart muscle.

Summation of Stimuli. If an isolated frog's ventricle which is not beating be stimulated with inadequate shocks, it may be found, on repeating these shocks at short intervals of time, that they culminate to cause a contraction of the ventricle. Stimuli which are subminimal may therefore summate.

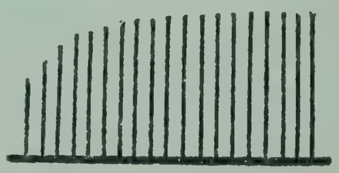


Fig. 348. Group of Contractions showing " Staircase " character.

Staircase. An improving effect of previous stimulation on the condition of heart muscle may be observed on the contractions themselves. Thus in a " Stannius preparation," if the ventricle be excited with single induction shocks once in every ten seconds, the first four or five contractions form an ascending series, each contraction being rather higher than the preceding one. This is often spoken of as the " staircase phenomenon " (Fig. 348). Each of the contractions represents a maximal contraction, but this maximum effort is at first not so great as it becomes after a few previous contractions. The phenomenon appears to be due to the accumulation of products of contraction which are beneficial.‡

The Refractory Period. When two stimuli are applied to any tissue at a sufficiently short interval of time, the second stimulus is without effect. In skeletal muscle the refractory period has passed away long before the muscle begins to respond, so that the stimulus applied after recovery from the refractory stage excites a second response, and wave-summation is produced.

In heart muscle the refractory period is extremely prolonged, *lasting as long as the contraction*, whether this occurs spontaneously or is evoked by a single stimulus. Hence, during the whole time that the muscle is contracted, any stimulus, however strong, applied to it remains without effect. Heart muscle therefore cannot be tetanised. As the contraction passes away the excitability once more rises and, as in the case of nerve, passes through a super-normal phase before it returns to normal. Directly after the contraction the excitability is so depressed that a very strong stimulus is required to produce excitation, but the strength necessary for this purpose gradually diminishes as the excitability returns. We may thus distinguish two phases : (*a*) an *absolute* refractory period during which no stimulus, however strong, can produce excitation ; (*b*) *a relative* refractory period, during which a second contraction is evoked by making the stimulus strong enough. The relation of these two phases in the excitability of cardiac muscle to the contraction produced by a single stimulus is shown in the accompanying diagram (Fig. 349). Directly after an effective stimulus is applied, the excitability of the preparation drops to zero and remains at this level during the phase of contraction. As relaxation of the cardiac muscle begins, excitability commences to rise, but some time elapses before it reaches normal. During this relative refractory period a second contraction of the muscle can be induced, the stimulus required becoming less as the absolute refractory period is left behind.

---

* Demoor. *Arch. Internat. de Physiol.*, 1924, **23**, 121 ; Demoor and Rylant *Com. rend. Soc. Belge. de biol.*, 1925, **2**, 814 ; *Arch. Internat. de Physiol.*, 1930, **32**, 80.

† Bowditch. Ber. Säcsh Ges., 1871, **23**, 652.

‡ Alison Dale. *Physiol.*, 1932, **75**, 1.

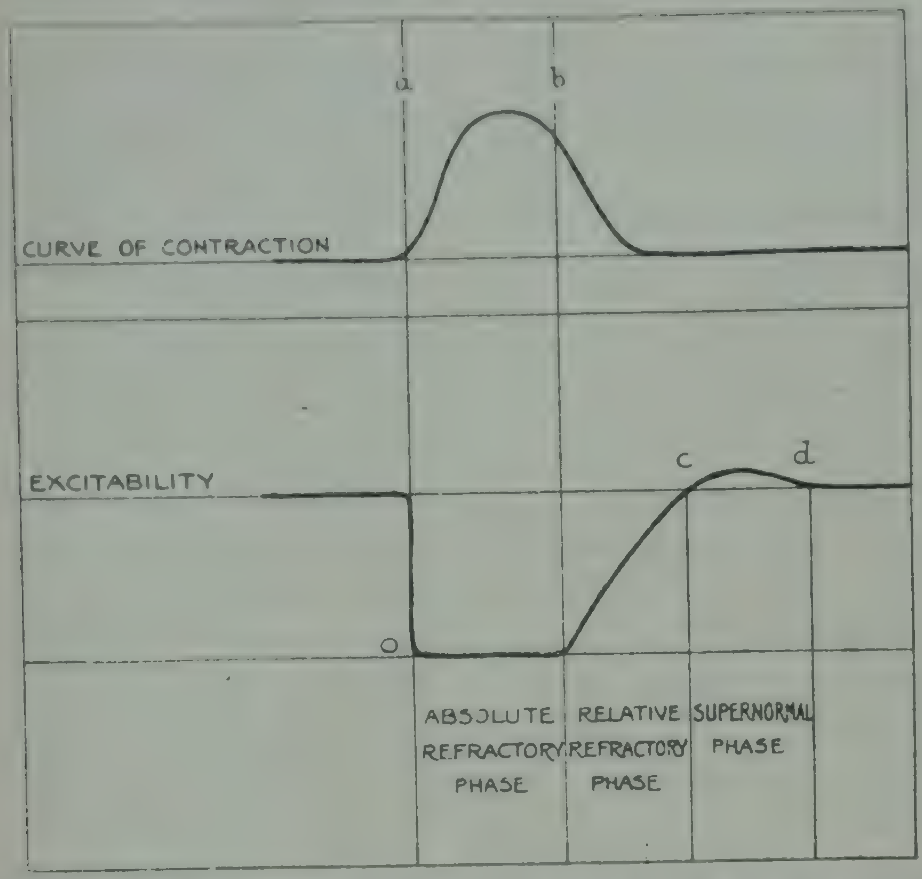


Fig. 349. Diagram showing changes in the excitability of heart muscle accompanying and following a single contraction.

A chemical explanation would seem to be applicable to these facts. Heart muscle, like skeletal muscle, contains adenosine tri-phosphate and creatine phosphate. As with skeletal muscle, part of these are broken down when contraction occurs, and probably provide the energy for the contraction.* After the contraction the material is resynthesised. We may suppose that, just after the sudden breakdown of the labile substances has taken place, any further breakdown becomes for a time impossible ; this time would be that of the absolute refractory period. As the store of labile material is regenerated, however, the excitability and contractility begin to return again, so that the tissue can respond to feebler and feebler stimuli, and at a certain point this store is great enough to enable the heart to respond to its normal " spontaneous " stimulus to contract. Just before this spontaneous contraction the heart muscle would respond to a minimal stimulus.

We should therefore expect the length of the refractory period to depend on the duration of the excitatory process : the smaller the breakdown as the result of excitation, the shorter will be the subsequent absolute refractory period. A contraction excited during the relative refractory period will be shorter than normal, because there has been less time for the accumulation of irritable material. Hence if the heart is stimulated to beat more frequently, the absolute refractory period following each beat is shortened, and we find that the length of the refractory period is inversely proportional to the frequency of the beat.

Thus Lewis found the following duration of the absolute refractory phase in mammalian auricular muscle under different rates of excitation :—

Rate per sec.	Refractory phase duration.
100	0·2 sec.
130–140	0·15 to 0·17 sec.
290	0·08 to 0·11 sec.

The refractory period is also shortened by stimulation of the vagus. This also will have the effect of diminishing the extent of the discharge at each beat. Thus the length of the refractory period seems to be determined by the amount of energy liberated or by the extent of the breakdown of irritable material in the previous contraction.

Under certain conditions the refractory period appears to be abnormally prolonged owing to the fact that a contraction is initiated by the second stimulus, but does not spread to the whole of the muscle, and hence is concealed (Drury and Love). This increase of the *effective* refractory period is seen when the muscle is subjected to cold, pressure, acid solutions, veratrine or quinidine, some of which, *e.g.* veratrine, do not alter the absolute refractory period at all, while others, *e.g.* quinidine, increase the absolute refractory period, but not to such an extent as might be supposed from estimations of the effective refractory period.

---

* Clark, Eggleton and Eggleton. *J. Physiol.*, 1932, **75**, 332.

Premature Contractions. When a ventricle is beating at its normal rhythm in consequence of the stimuli communicated to it from the sinus venosus, the application of a single shock to the ventricle early in its diastole, is followed by the appearance of a ventricular contraction some while before the normal one would be due. It is called a *premature contraction* (or, less accurately, an " extra systole "), and is followed by a " compensatory pause " (Fig. 350). This is due to the fact that one of the impulses travelling from the sinus arrives at the ventricle during the refractory period ensuing on the application of the artificial stimulus ; hence it produces no effect and the ventricle has to wait for the arrival of the next succeeding excitatory wave from the sinus before it gives its next beat. In certain cases the first contraction after the pause is considerably augmented. The compensatory pause, of course, does not occur when we are testing the effects of artificial stimuli on the sinus venosus.

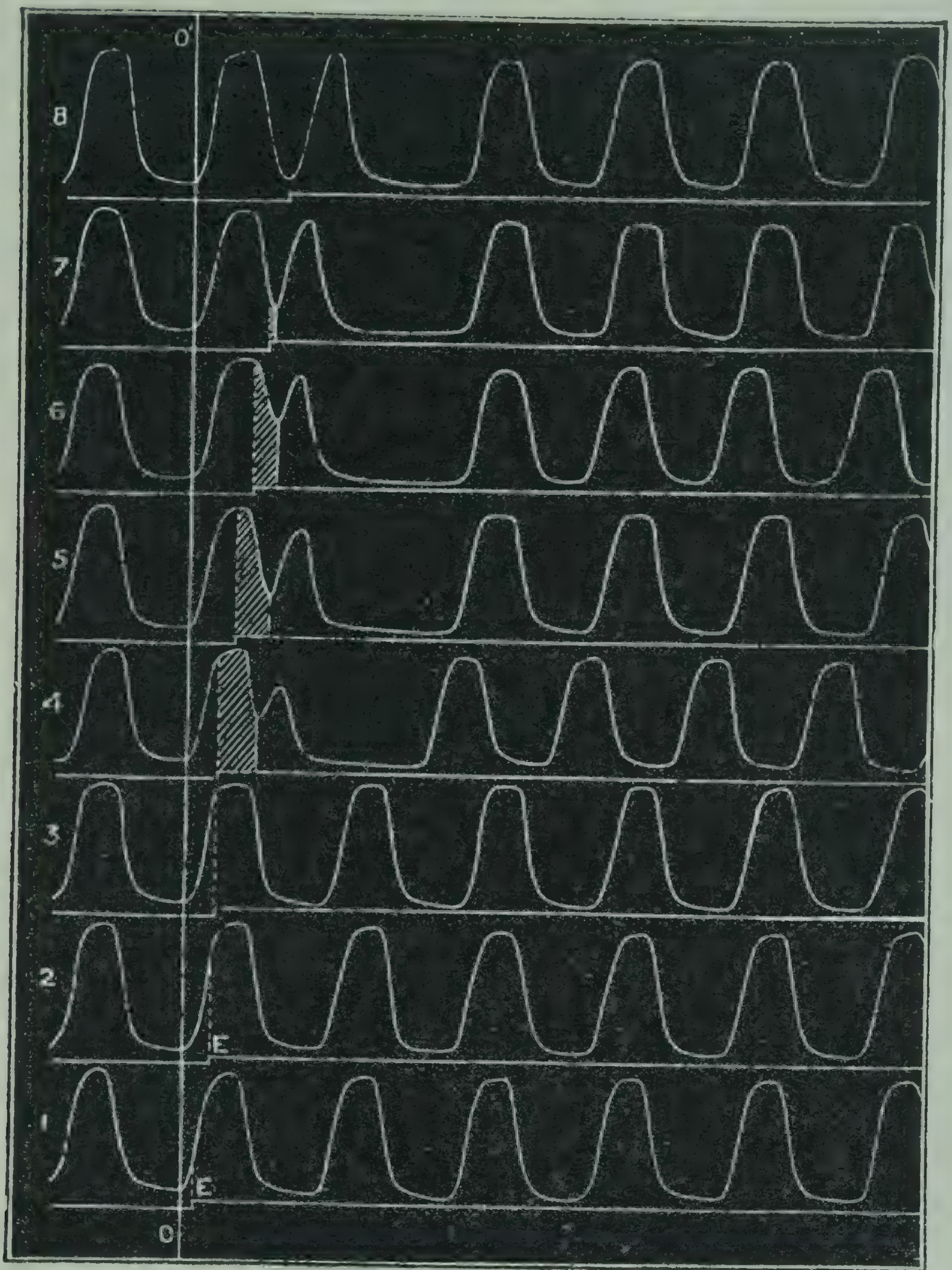


Fig. 350. Tracings of Spontaneous Contractions of Frog's Ventricle, to show Refractory Period. (Marey, *La méthode graphique,* 1885, p. 417 (Paris, Masson).)

In each series the ventricle was stimulated by an induction shock at E, as indicated by the tracing of the signal. In 1, 2 and 3, this stimulus had absolutely no effect, since it fell during the refractory period. In 4, 5, 6, 7 the effect of the shock was to interpolate a premature contraction in the series, the latent period (shaded part) gradually diminishing from 4 to 7. In 8 the irritability of the preparation was already considerable, and the latent period inappreciable. The " compensatory pause " after the premature contraction is also well shown in 4, 5, 6, 7, 8.

INVERTEBRATE HEARTS. The hearts of invertebrates are often very different in structure and properties from those of the higher animals. In *Limulus*, which has been extensively studied by Carlson,* the heart forms a segmented tube of ordinary striated muscular fibres, which has a local system of ganglion cells, so situated that they can be cut away entirely from the muscular portions of the organ. Extirpation of the nerve cord abolishes spontaneous contractions. It is possible to make a nerve-muscle preparation of the anterior part of the heart, consisting of the muscle of the first two segments with a longer stretch of the lateral nerves. Stimulation of the lateral nerves with a single shock causes a single beat of the anterior segments; tetanising shocks cause a tetanus. The beat of the heart is originated and co-ordinated by the action of the local ganglionic centres. It would, for these and similar reasons, be wrong to transfer the results obtained on the heart of Limulus, or other invertebrates, to the explanation of the phenomena exhibited by the hearts of vertebrates.

THE MAMMALIAN HEART

In the mammal each side of the heart consists of an auricle into which the veins open, and a ventricle which receives the blood from the auricle and discharges it into the arterial trunk—either aorta or pulmonary artery. Since the auricles have to act merely as a receptacle for *part* of the blood which enters during the relaxation or diastole of the heart, their cavities are smaller than those of the ventricles, and their walls are thin, corresponding to the small amount of work thrown on them in propelling blood into the relaxed ventricle. The ventricles have the office of carrying on the main work of the circulation, viz. of forcing blood through the peripheral resistance. Their walls are much thicker than those of the auricles. The right ventricle has a wall which is only

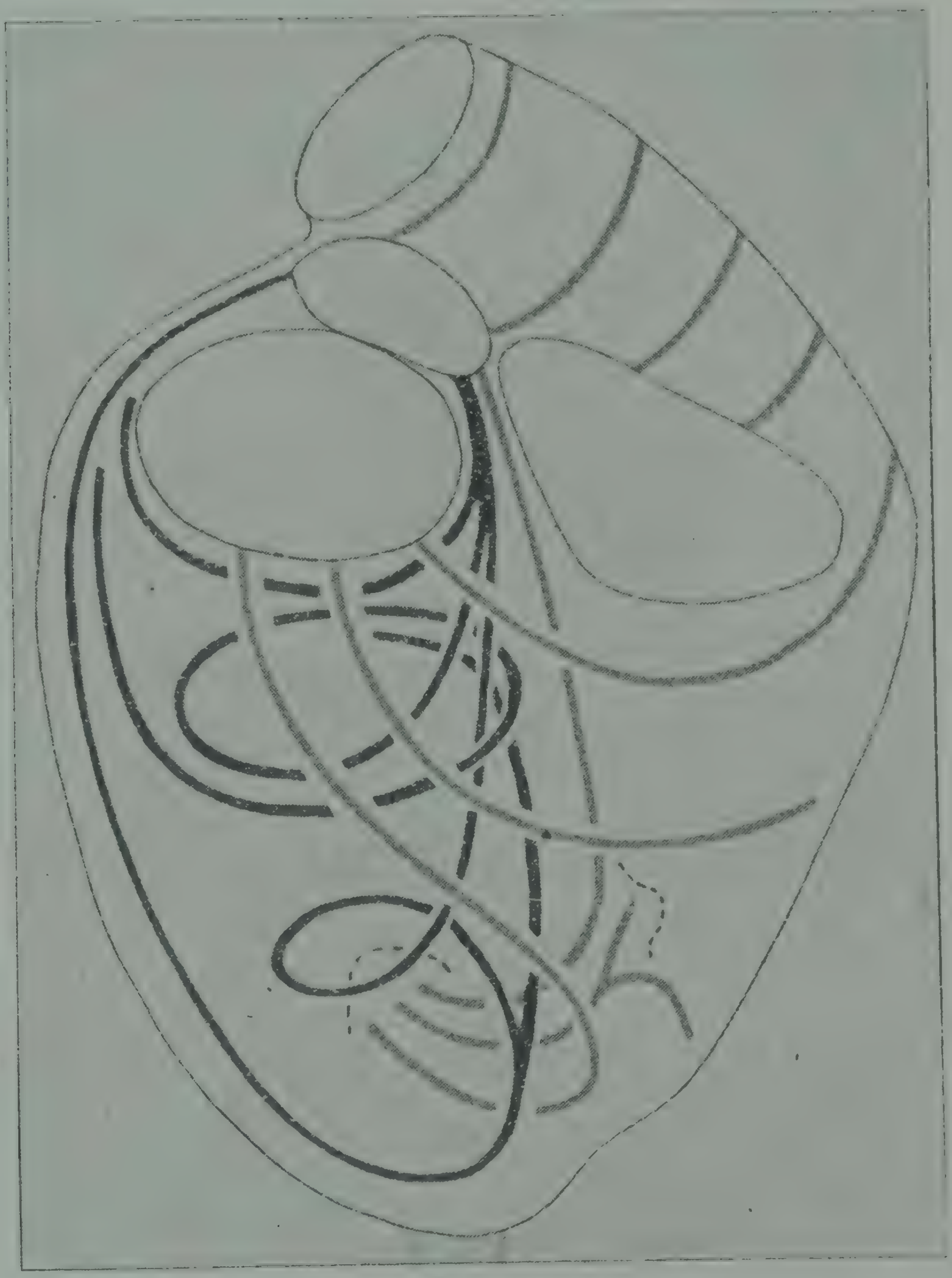


Fig. 351. View of the Heart from behind, to show the Course of the Chief Strands of Muscle Fibres. (Mall.)

The black lines represent the *bulbo-spiral* fibres, the grey lines the *sino-spiral* fibres.

---

* Carlson. *Amer. J. Physiol.*, 1904, **12**, 67.

about one-fourth the thickness of the left ventricle, in conformity with the much heavier work to be done by the latter. The capacity of both ventricles is approximately equal.

The auricles are separated from the ventricles by a fibro-tendinous ring. From this ring take origin most of the muscular fibres of the heart walls.

The muscular fibres of the auricles run in both circular and longitudinal directions (Fig. 351), the circular fibres being continued round both auricles and special rings of circular fibres surrounding the openings of the great veins. From the fibro-tendinous ring between the auricle and the left ventricle and from the sides of the aorta, the muscular fibres forming the superficial layer of the ventricular wall pass obliquely downwards to the left towards the apex of the ventricle. Here they loop round into the interior of the ventricle and pass up near its inner surface to end either in the papillary muscles or in the auriculo-ventricular ring of fibrous tissue. Between these two layers we find a third median layer of muscular fibres which is in the form of a muscular cone. The fibres of this layer form complete loops round the left ventricle. The middle layer is connected by many strands of muscular fibres with both inner and outer layers.*

As Harvey showed, the normal direction of the blood flow through the heart is determined mainly by the valves. The auriculo-ventricular valves are membranes attached round the entire circumference of the auriculo-ventricular ring (Fig. 352). They are composed of fibrous and elastic tissue covered on each

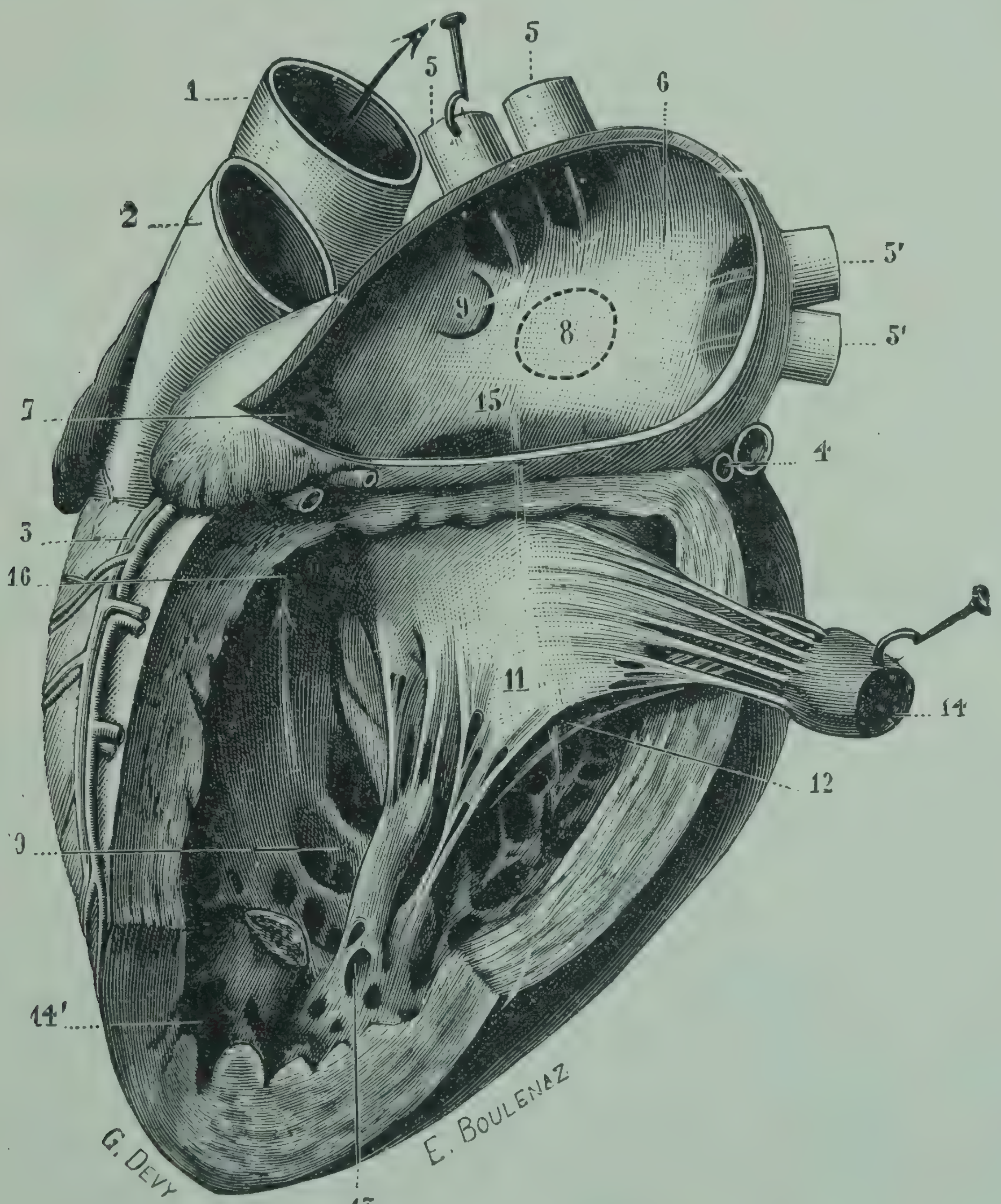


Fig. 352. Left Auricle and Ventricle, with Outer Side cut away to show Chief Points in Anatomy of Heart. (Testut.)

1. Aorta.
2. Pulmonary artery.
3. Ant. coronary vessels.
5, 5′. Pulmonary veins.
6. Left auricle.
7. Auricular appendage.
10. Cavity of left ventricle.
11, 12. Mitral valves.
13, 14. Papillary muscles.
16. Arrow pointing to aortic orifice.

---

* *v.* Lower. *Tractatus de Corde*, London, 1669, p. 24. (Facsimile and translation by Franklin, Oxford, 1932.)

side with endocardium, and project downwards into the cavities of the ventricles. On each side the membrane is divided by deep incisions into large flaps, three in number on the right side (the tricuspid valve) and two in number on the left side (the mitral valve). The sail-like margins of these valves are connected by thin tendinous cords to the papillary muscles, which are projections of the muscular walls of the ventricles. By this means the edges of the valves are kept close together and prevented from eversion under the strong pressure exerted by the blood in the contracting ventricle. By the downward pull of the papillary muscles on the valves during the contraction of the ventricles, closure is rendered more complete, the inner surface of the valves being apposed over a considerable area. The action of the valves is aided by the contraction of the fibres surrounding the base of the heart, so that the auriculo-ventricular orifice is much smaller during systole than during diastole.

Each arterial orifice is provided with three valves, semilunar in shape and each attached by its convex borders to the arterial wall, and presenting in the middle of its free border a small nodule, the *corpus Arantii*, from which fine elastic fibres pass to all parts of the valve. The extreme margin of the valve, the *lunula* on each side of the corpus Arantii, is very thin. Whenever the pressure in the arteries is greater than that in the ventricles, these valves are closed, and the thin margins come in contact with similar portions of the adjacent valves, so preventing any reflux of blood. The borders of the valves under these circumstances come together in the form of a star composed of three lines at angles of 120°, the three corpora Arantii being pressed together at the centre of the star.

No valves are found at the orifices of the great veins into the auricles, and so a reflux of blood occurs during contraction of the heart, but it is limited by the contraction of the muscular rings round the veins, which slightly precedes the main auricular contraction.

The heart and the roots of the great vessels lie almost free inside the pericardium. This is attached below to the central tendon of the diaphragm, and above to the arterial trunks. It is lined by a layer of mesothelium continuous with a similar layer covering the surface of the heart. The two surfaces are kept continually moist by the pericardial fluid, so that the heart can move freely within the pericardium without friction. One of the chief functions of the pericardium appears to be to check an excessive dilatation of the heart.

THE CONTRACTION OF THE MAMMALIAN HEART

A perfectly rhythmic activity of the mammalian heart may be maintained for hours after its removal from the body by perfusing the coronary arteries with blood, or with a suitable warm oxygenated perfusion fluid.

It is possible to sever all the nervous and muscular connections between auricles and ventricles without stopping the contractions of the ventricles which, after a pause, recommence beating at a rhythm which is independent of and slower than that of the auricles. A piece of an auricular appendix, suspended in warm oxygenated Ringer's solution, will continue to beat for hours, and a fragment of the ventricular wall, free from ganglion cells, will similarly contract if fed by an artificial circulation through a branch of the coronary artery. We may therefore conclude that in the mammalian, as in the amphibian, heart every part of the heart muscle possesses the power of rhythmic activity, the normal sequence of transmission being determined by the greater frequency of the natural rhythm at the venous end of the heart.

Conduction of Excitatory Wave. The excitatory condition started at one point in the muscle spreads through it in all directions, and the spread is independent of nerve fibres. As in the frog, the excitatory process may, in exceptional circumstances, be made to travel in a direction the reverse of normal, as, for instance, when the ventricles are excited at a rhythm of

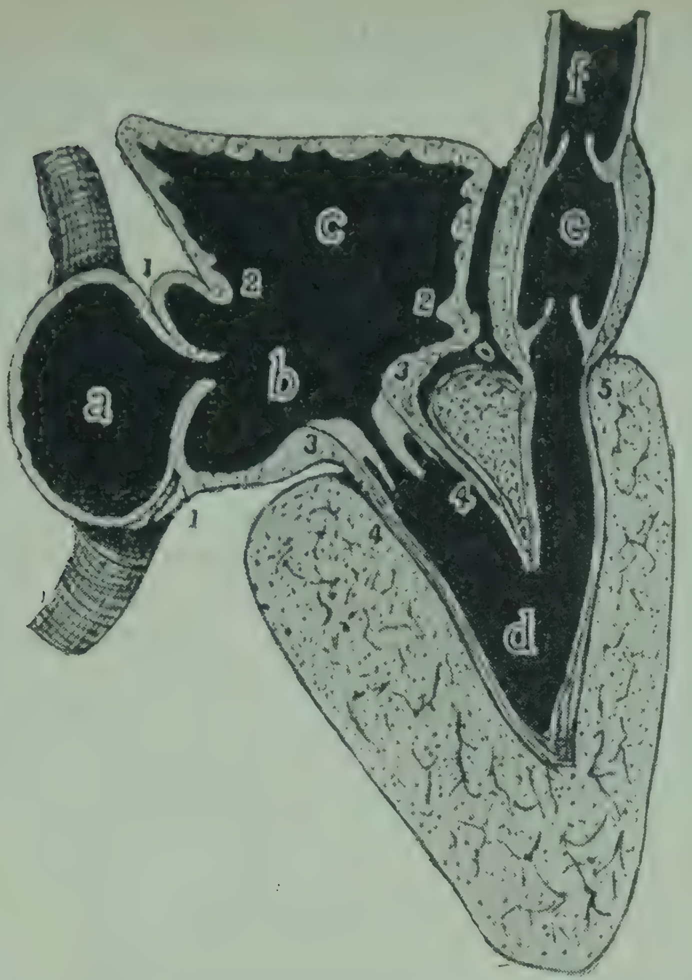

f
e
c
2
2
1
1
a
b
3
3
4
4
5
d

FIG. 353. A Generalised Type of Vertebrate Heart. (KEITH.)

a. Sinus venosus.
b. Auricular canal.
c. Auricle.
d. Ventricle.
e. Bulbus cordis.
f. Aorta.
1–1. Sino-auricular junction and venous valves.
2–2. Canalo-auricular junction.
3–3. Annular part of auricle.
4–4. Invaginated part of auricle.
5. Bulbo-ventricular junction.

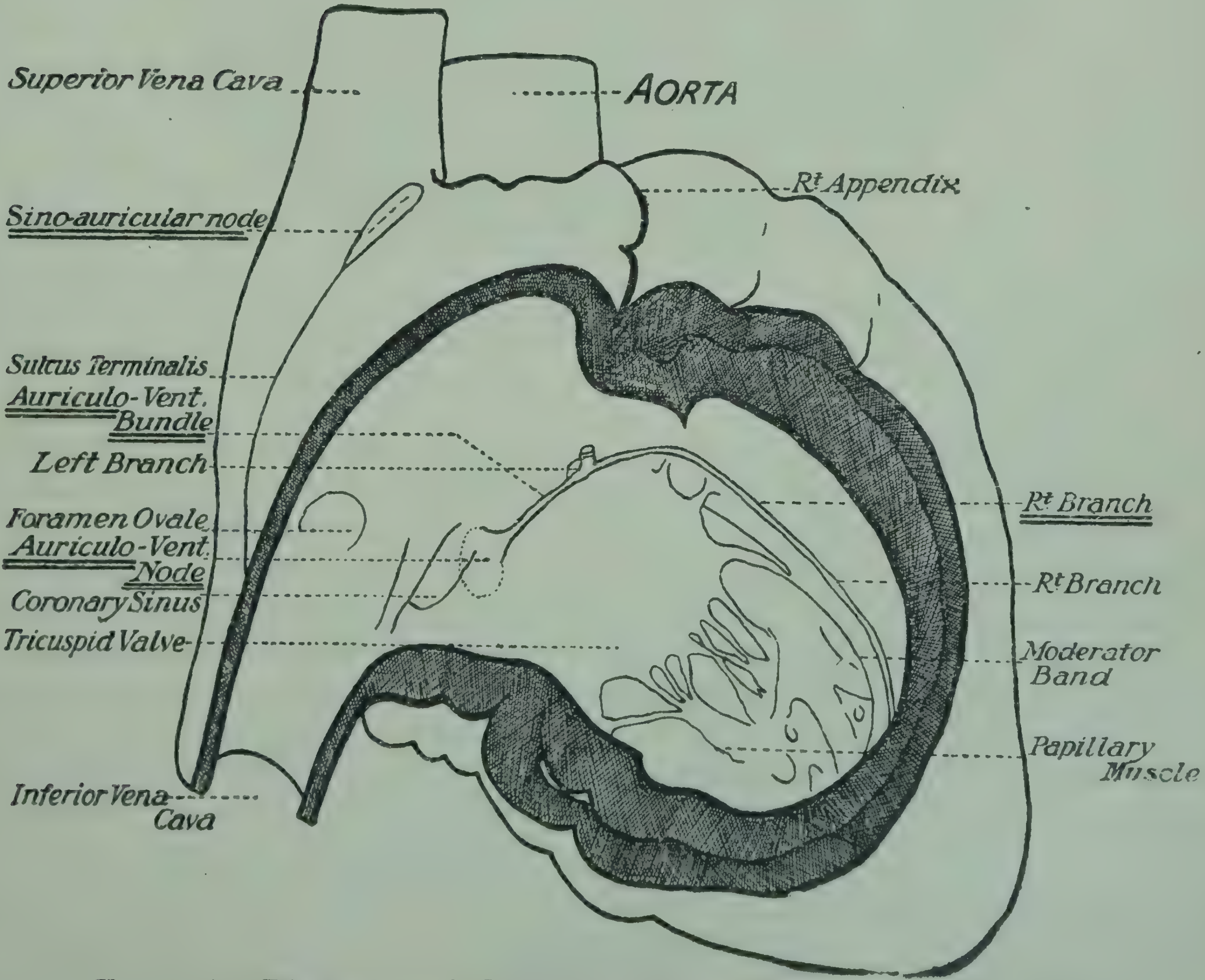

Superior Vena Cava
AORTA
Rt Appendix
Sino-auricular node
Sulcus Terminalis
Auriculo-Vent. Bundle
Left Branch
Rt Branch
Foramen Ovale
Auriculo-Vent. Node
Rt Branch
Coronary Sinus
Tricuspid Valve
Moderator Band
Papillary Muscle
Inferior Vena Cava

FIG. 354. Diagrammatic Representation of Course of A.V. Bundle.

higher frequency than the natural one starting at the venous end of the heart.* The phenomena are easily explained on the assumption that the whole of the musculature of the heart acts in many respects as a single muscle fibre, along which an excitatory process may be propagated in any direction. But in the mammalian heart the muscle fibres of both auricles and ventricles are seen to arise from a fibro-tendinous ring at the auriculo-ventricular junction, leaving apparently no muscular continuity between the two cavities. The discovery by Kent,† and then by His ‡ of a distinct band of modified muscle fibres passing, in the septum, from the auricles to the ventricles, the "*auriculo-ventricular bundle*," furnished an anatomical basis for the conduction of the wave.§

In the adult heart, *e.g.* of man, the anatomical relations of the different cavities have become considerably modified from their primitive form (Fig. 353) in the course of development. The sinus venosus, *i.e.* the part where in the lower vertebrates the contraction wave takes its origin, is now represented merely by the termination of the superior vena cava and of the coronary sinus in the right auricle, in a small amount of tissue underlying the *sulcus terminalis* of the right auricle, as well as by the remains of the Eustachian and venous valves. The auricular canal gives rise to the auricular septum and to the auricular ring surrounding the auriculo-ventricular orifice, and in some hearts it is prolonged into the ventricle as the intraventricular or invaginated part of the auricular canal.

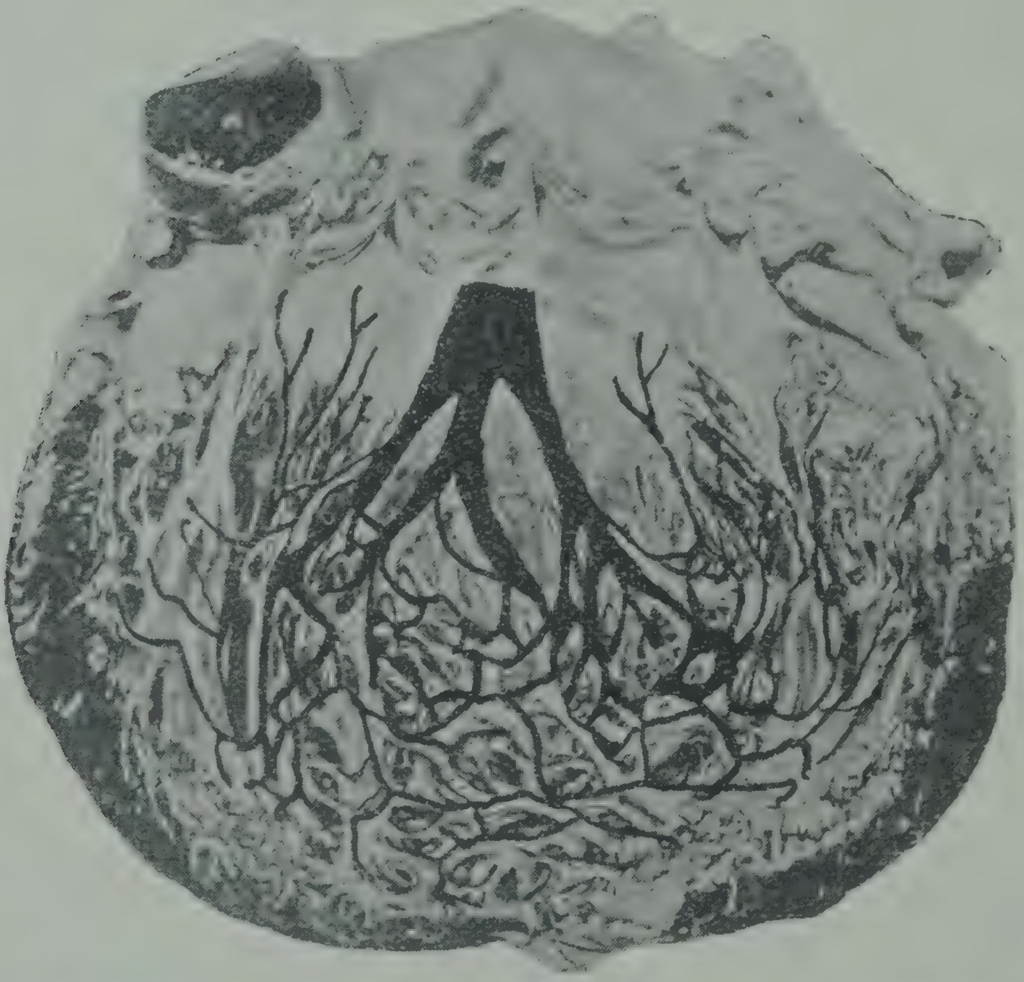


FIG. 355. Left Ventricle, laid open to display the Interventricular Septum, on which the Course of the Left Division of the Auriculo-ventricular Bundle and its Ramifications are shown in black. (After TAWARA.)

In the adult heart two accumulations of more primitive tissue are found in the region corresponding to the sinus venosus of the embryo, and these are known as the *sino-auricular node* ‖ and the *auriculo-ventricular node*. They are composed of delicate interlacing fusiform fibres, faintly striated and embedded in dense connective tissue. The fibres contain very little glycogen. The sino-auricular node (Fig. 354) lies in the groove between the superior vena cava and the right auricle.¶ The auriculo-ventricular node lies at the base of the auricular septum on the right side, below and to the right of the opening of the coronary sinus (Fig. 355).

---

* McWILLIAM. *J. Physiol.*, 1888, **9**, 167.

† KENT. *J. Physiol.*, 1892, **14**, xxiii; 1893, **14**, 233.

‡ HIS. *Arbeiten. aus der med. Klin.*, 1893, **14** (Leipzig).

§ TAWARA. "Das Reizleitungssystem des Säugetierherzens," Jena, 1908; TODD. *Johns Hopkins Hosp. Rep.*, 1919, **19**, 1.

‖ KEITH and FLACK. *J. Anat. and Physiol.*, 1907, **41**, 172.

¶ BLAIR and DAVIES. *J. Anat.* 1935, **69**, 303.

Continuous with the A–V node, a bundle of muscular fibres (the auriculo-ventricular or A.V. bundle) runs along the interventricular septum and then divides into the right and left septal divisions, which pass down in each ventricle on the interventricular septum. Each division gives off branches which break up, finally forming a reticulated sheet of tissue over the greater part of the interior of the ventricles just below the endocardium, with branches that penetrate to end beneath the epicardium, and in the papillary muscles. (Figs. 354 and 355). The fibres composing this tissue are more primitive in character than the rest of the cardiac musculature and have long been distinguished as the "fibres of Purkinje." In them the striation is confined to the periphery of the muscle cell (Fig. 356). They are distinguished by a high glycogen content. The course of the system of Purkinje tissue in the heart may be demonstrated by the injection of coloured materials into the sheaths which enclose the main strands.*

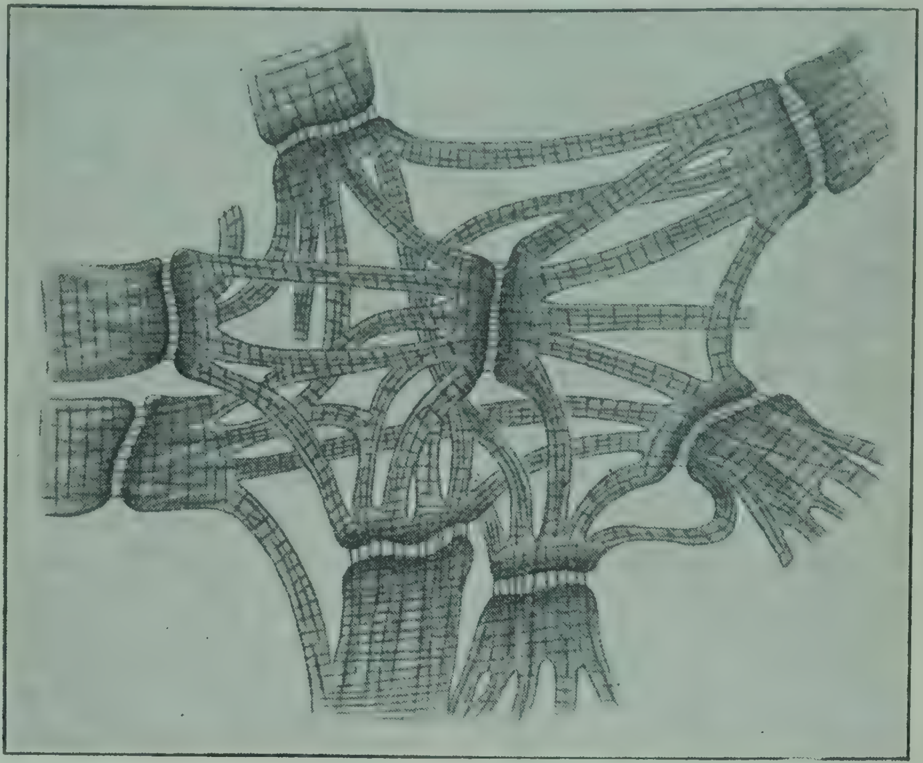


Fig. 356. Fibres of Purkinje, from the Sub-endocardial Network. (Tawara.)

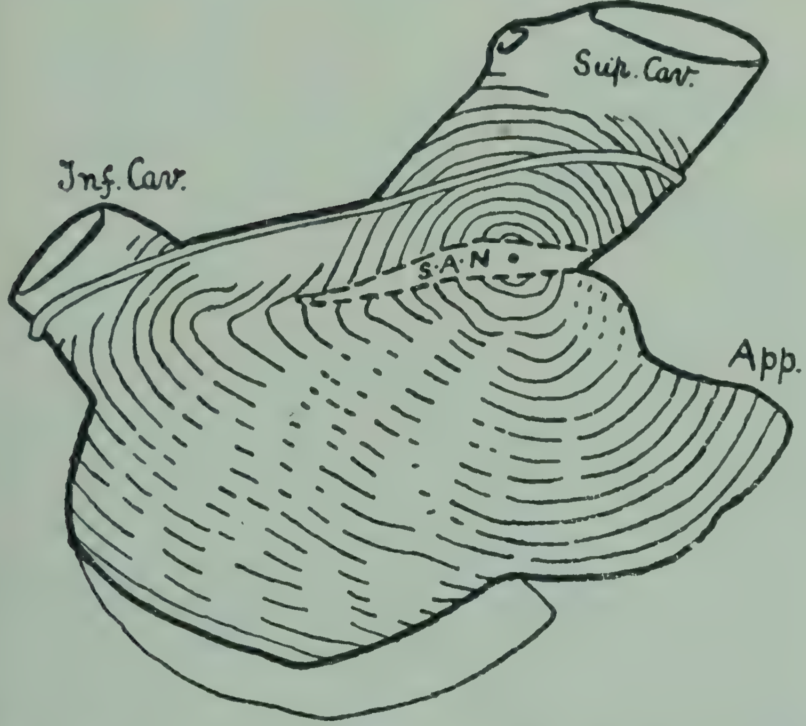


Fig. 357. Diagram showing the manner in which the Excitatory Process spreads from the Sino-auricular Node over the Auricles and along the great Veins. (T. Lewis.)

There are plexuses of nerves on the outer surface of the heart and beneath the endocardium, and ganglion cells around the openings of the great veins, along the interauricular septum, in the auriculo-ventricular groove and in the basal parts of the ventricles. Numerous nerve fibres and ganglion cells accompany the muscle fibres of the auriculo-ventricular bundle. We have, however, no reasons for regarding the nervous structures as primarily concerned in the initiation or propagation of the excitatory wave.

The auriculo-ventricular bundle forms the only continuous muscular tissue between the auricles and ventricles, and destruction of it causes complete abolition of the normal sequence of beat between auricles and ventricles.

By leading off different parts of the heart to the string galvanometer, it is possible to follow the progress of the excitatory process. It is then found that the sino-auricular node is normally the starting-point of the excitatory process concerned in each heart beat. It is therefore spoken of as the "*pace-maker*" of the heart. At each beat an excitation starts at the sino-auricular node, spreads a short way up the great veins, and along the auricular muscle in all directions at a speed of about 1 m.p.sec. (Fig. 357). When it arrives at the auriculo-ventricular node, this in turn is excited; it conducts slowly (0·2 m.p.sec.) and so there is some delay in transmission here. The impulse is then carried on to the ventricles along the auriculo-ventricular bundle, spreading along the branches of this bundle to almost all parts of the ventricular muscle. Although normally the sino-auricular node initiates each heart beat, this node can be put out of action by injury or by cooling without stopping the rhythmic sequence of the heart beat, the office of pace-maker being now taken up by the auriculo-ventricular node. A specialisation of function accompanies the differentiation in structure which we find in the auriculo-ventricular bundle and its branches. The conduction of the excitatory process along the auriculo-ventricular bundle, *i.e.* Purkinje tissue, occurs about ten times as fast as the conduction through the ordinary muscular tissue of the heart, the rates being about 5 m. and 1 m. per second respectively.

---

* Lhamon. *Amer. J. Anat.*, 1912, **13**, 55; Cohn. *Heart*, 1912, **4**, 225.

The limitation of the muscular continuity to a single narrow bundle, which is endowed with greatly increased conducting powers and ends in a network of similar tissue, ensures that all parts of the ventricles contract almost simultaneously. If this were not the case, the sudden contraction of the muscle fibres near the base of the ventricles would simply bulge out the still uncontracted portion near the apex, and there would be a risk of injury to the uncontracted part of the ventricle.

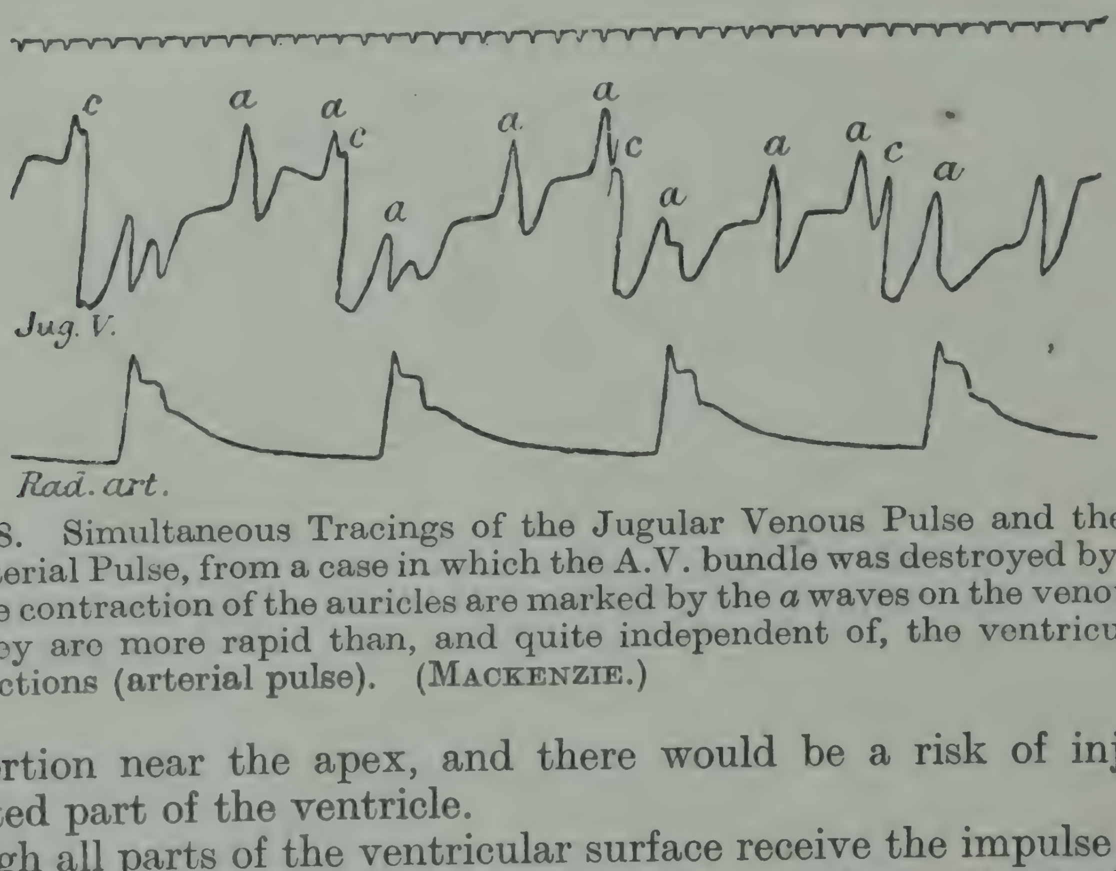


FIG. 358. Simultaneous Tracings of the Jugular Venous Pulse and the Radial Arterial Pulse, from a case in which the A.V. bundle was destroyed by disease. The contraction of the auricles are marked by the *a* waves on the venous pulse. They are more rapid than, and quite independent of, the ventricular contractions (arterial pulse). (MACKENZIE.)

Although all parts of the ventricular surface receive the impulse to contract almost simultaneously, *i.e.* within 8 msec. of one another in the dog's heart, the whole ventricle being involved within 0·07 sec., the contraction wave is found to commence 10–17 msec. earlier at the surface overlying the interventricular septum. On the other hand, the conus of the pulmonary artery is the last part of the heart to begin contracting.

The conductivity of the A.V. node and bundle, the former never high, may be impaired by disease. In such cases we get a series of phenomena known as *heart block*, the main characteristic of which is the slow rhythm of the ventricle, accompanied by a normally rapid auricular rhythm, seen in the venous pulse to be entirely independent of the ventricular pulse. The automatic activities of auricle and ventricle are, in fact, dissociated. At certain intervals or at earlier stages of the disease, the fibres of the bundle may present only a partial block, so that the ventricle responds once to every second contraction of the auricle. The existence of this condition is shown at once on the electrocardiogram (or venous pulse, see Fig. 358) by the dissociation of the normal relation between the auricular and ventricular components.

Fibrillation. If a strong faradic stimulus is applied during the latter part of systole to the ventricle of a heart beating normally, a premature contraction results and then the rhythmic contraction of the chamber stimulated sometimes ceases instantaneously; the chamber dilates and its walls are seen to be the seat of vermicular movements, each muscle fibre or fasciculus seeming to be contracting independently of the fibres around it, so that the whole surface is affected with a writhing movement.* The same condition of fibrillation may sometimes be produced in the mammalian ventricle by obstruction of one coronary artery, and may account for the sudden death which frequently occurs as a result of occlusive disease of the coronary arteries. It seems probable that a similar condition is sometimes responsible for the sudden death which may occur during the induction of chloroform anæsthesia.†

Fibrillation in the ventricles is usually irreversible and causes death.

When fibrillation is induced in the auricles, it often passes off after a variable time and the ordinary rhythmic contractions recommence. Fibrillation of the auricles is a frequent complication of heart disease, where the myocardium has been injured by some previous infective process, such as rheumatic fever. The fibrillation is at the rate of about 450 per minute, and does not spread to the ventricles, but sets up frequent and irregular excitations which travel along the A.V. bundle and cause a rapid and irregular ventricular rhythm. The smaller contractions in this rhythm may be ineffective for the expulsion of blood from the ventricles, and merely promote the exhaustion of the heart.

In some instances the auricular fibrillation is not general, and the effect of the local fibrillation is to impart a very rapid, but regular, rhythm to the auricles, while the ventricles have a regular rate one-half or a quarter that of the auricles. This is called *auricular flutter*.

The clue to the causation of fibrillation was given by observations of Mines ‡ and of Garrey. Mines found that under certain conditions a ring of frog's ventricular muscle might respond to a single stimulus by a wave of contraction passing continuously round and round the ring, and Garrey observed the same phenomenon in a ring cut from the contractile tissue of the umbrella of the jelly fish. The way in which this circus movement is brought about has been elucidated by T. Lewis. In Fig. 359 is represented a ring of auricular tissue. If it is excited rhythmically at X, the wave of excitation will travel in both directions round the ring and finish at Y. If, however, the rate of stimulation be continually increased, the refractory period gets steadily shorter, and a time will arrive at which any slight difference in the duration of the refractory stage at *a* and *b* becomes important. Let us suppose it is shorter at *b*. The wave due to the second stimulus may therefore be able to pass *b* but not *a*. The wave will thus pass in one direction only (anti-clockwise), and will travel right round the ring. By the time it has travelled past Y and arrived at *a*, the refractory stage may have passed off, so that the muscle at *a* can now contract. Whether the circus movement continues indefinitely will depend on (1) the rate of conduction, (2) the length of the ring, and (3) the duration of the refractory period. Anything which diminishes the rate of conduction and the duration of the refractory period (and both these changes result from frequent stimulation) will tend to set up a circus movement. When fibrillation is present, there is a circus movement in a ring of tissue round the mouths of the great veins. This ring is not, however, separated from the rest of the auricle, so that the wave of excitation in its course round the ring will excite in turn the adjacent portions of the auricular muscle, so that these will contract at all intervals after the passage of the wave past X. All co-ordinated movements will then be abolished and the auricle will continue in a state of fibrillation. This explanation is supported by the fact that it is possible to set the auricle into fibrillation by the application of two shocks at such an interval that the second falls just outside the absolute refractory period of the auricular muscle, and the same thing doubtless occurs when the auricle or ventricle is set

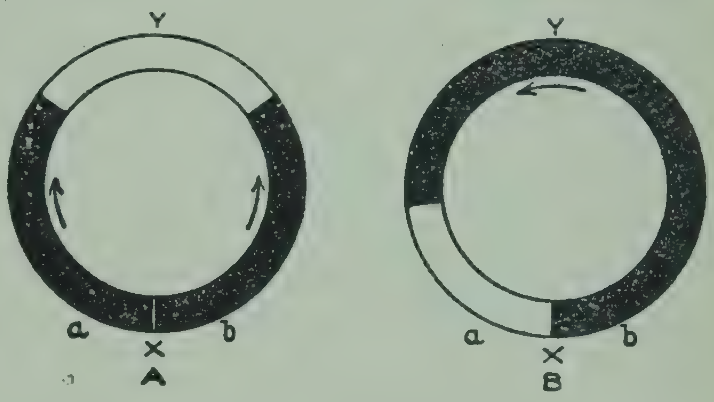


Fig. 359.

---

* Wiggers and Wégria. *Amer. J. Physiol.*, 1940, **128**, 500.

† Levy. *J. Physiol.*, 1911, **42**, Proc. iii ; **43**, Proc. xviii.

‡ Mines. *J. Physiol.*, 1913, **46**, 349.

into fibrillation by the application of strong faradic currents. Administration of quinidine may put a stop to auricular fibrillation owing to the fact that it increases the effective refractory period.

For Reference

Eyster and Meek (1921). The Origin and Conduction of the Heart Beat. *Physiol. Rev.*, **1**, 1.

Garrey (1924). Auricular Fibrillation. *Physiol. Rev.*, **4**, 215.

Lewis (1925). "The Mechanism and Graphic Registration of the Heart Beat." (London, Shaw.)

Rijlant (1931). "La conduction dans le cœur." (Liège, Waillant-Carmanne.)

THE SEQUENCE OF EVENTS IN THE CARDIAC CYCLE

The heart beats about seventy times a minute in normal resting human adults. Each beat begins by a simultaneous contraction of both auricles, and of the auricular appendages, which become pale and bloodless. The contraction of the auricles is followed by that of the ventricles, as a result of which blood is thrown out into the large arteries. In clinical terminology, the term "systolic" refers to the ventricular systole, so that "*pre-systolic*" events occur during auricular systole. The duration of ventricular systole in man is given by the interval between the commencement of the first and second heart sounds, or by the interval between the upstroke of the R wave and end of T wave on the electrocardiogram. The contraction of the auricle lasts about a tenth of a second, that of the ventricles about three-tenths of a second. The period of total relaxation or diastole lasts about four-tenths of a second. During this cycle of changes the following events are taking place within the heart :

In the diastolic period the aortic valves are closed, and the auriculo-ventricular valves open. There is a continuous flow of blood from veins, through auricles into ventricles and, as the walls of both these cavities are relaxed, there is no impediment to the inflow of the blood until the heart is quite full and begins to stretch the pericardium. Under normal circumstances the diastole comes to an end before such restraining influence of the pericardium can be effective.

The contraction of the auricles drives some of their contents into the ventricles and so still further increases their filling, no resistance being offered by the flaccid wall of the ventricles. The amount of blood propelled into the ventricle by the auricular contraction has been variously estimated as from 18 to 60 per cent. of the total ventricular filling. In fact it is a variable quantity, depending on complex dynamical factors, chief of which is the time relation between auricular and ventricular systoles.* But even when the auricle is fibrillating, and non-functional, ventricular filling may still be adequate. As the blood rushes from auricle into ventricle through the funnel-shaped opening of the membranous tube formed by the valves, eddies are set up in the ventricle tending to close the valves, so that they are held, by the resultant of the two opposing currents, in a condition midway between closure and opening.

The onset of the ventricular contraction is extremely sudden. There is a quick rise of pressure in the ventricle, which presses together the flaps of the mitral or tricuspid valve, while the bases of these valves are approximated by the contraction of the circular fibres at the base of the ventricles. As the heart shortens in systole the papillary muscles also shorten, so that the valves are prevented from being everted into the auricles, while the blood is pressed, so to speak, between the cone of the ventricular wall and the cone formed by the closed valves.

The outflow of blood from the ventricles, however, does not commence instantly, and for a very short time, therefore, the ventricle is contracting on a closed cavity, and is thus performing an *isometric* contraction (isometric phase), during which the pressure within it must rise abruptly. Before the beginning of systole the pressure in the ventricular cavity is quite small (only 2 or 3 mm. Hg), while the pressure in the aorta is 70 to 90 mm. Hg. Before the aortic valves can be opened, the pressure in the left ventricle must rise to a point which is greater than that in the aorta, and similarly on the right side of the heart. As soon as this happens the valves open and the outflow of blood (ejection phase) commences, and continues so long as the pressure in the ventricles is higher than that in the great arteries.

---

* Jochim. *Amer. J. Physiol.*, 1938, **122** 639.

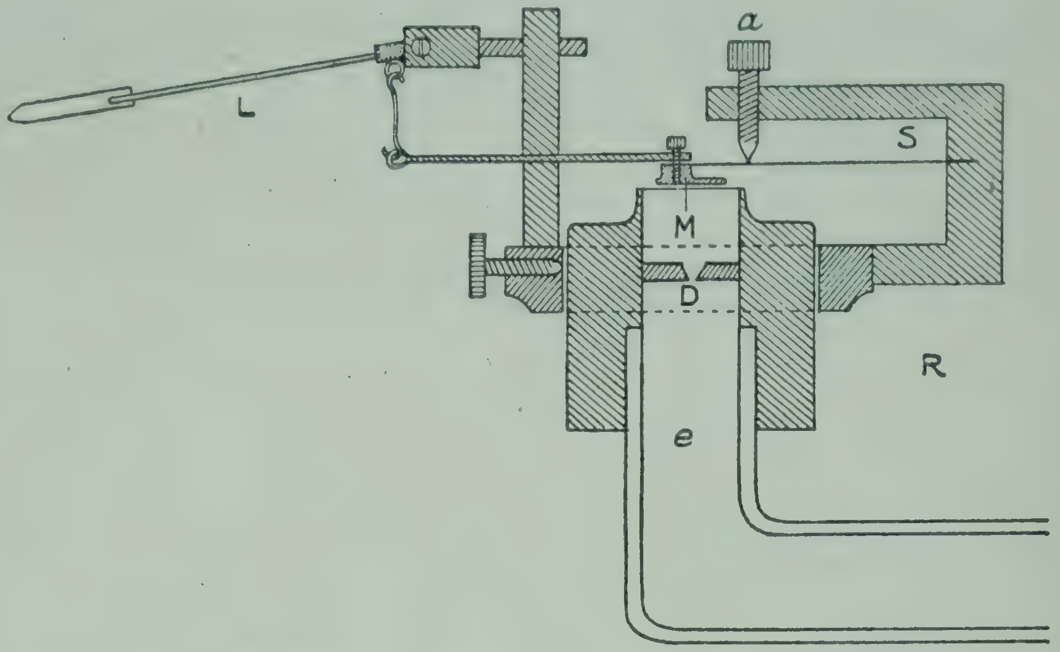


FIG. 360. Diagram to show the principle of the membrane manometer. M, the rubber membrane; S, the spring; *e*, connecting tube to artery; D, damping diaphragm; L, the lever. (FRANK.)

Even in extreme contraction the cavities of the ventricles are never entirely obliterated, though the right ventricle is reduced to a narrow slit widening out slightly in the neighbourhood of the auriculo-ventricular orifices, while in the left ventricle a distinct cavity is left between the mitral valves and the free ends of the papillary muscles. During normal activity the emptying of the cavities rarely proceeds to so great an extent. Directly the ventricular pressure falls below the arterial pressure, the arterial valves must close again and the output of blood come to an end.

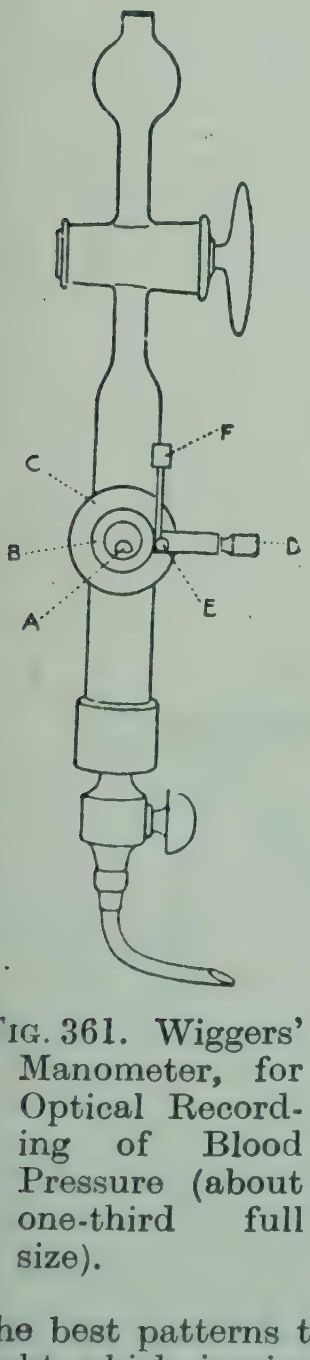


FIG. 361. Wiggers' Manometer, for Optical Recording of Blood Pressure (about one-third full size).

In order to obtain an accurate idea of the progress and duration of each of these events in the cardiac cycle, it is necessary to study the changes occurring in the pressure within the auricles and ventricles during the various phases of the heart beat.

The Endocardiac Pressure. A manometer which shall register accurately the changes in the pressure within the heart must be capable of responding faithfully to very rapid changes without entering into oscillations of its own. Thus in the left ventricle at the beginning of the systole, there may be a rise of 130 mm. Hg in 0·06 sec., *i.e.* at the rate of nearly 3 atmospheres per sec. In a heart beating rapidly and forcibly, the rise may be as rapid as 8 atm./sec. A mercurial manometer would be quite unsuitable.

A study of the theory of recording instruments by Otto Frank * showed that the fundamental requirements of such an instrument were that the mass of fluid and of parts of the instrument moved must be as small as possible: the ratio $\Delta P/\Delta V$, of pressure change per unit of volume change, must be high; also, in order to avoid vibrations of instrumental origin, the period of self-vibration of the mass moved must be as high as possible. This is accomplished, as in Hürthle's instrument, by using a very small tambour, covered with a strong, tightly stretched membrane connected, by as short and wide a tube as is feasible, with the heart or blood vessel where it is desired to register changes of pressure. The whole is filled with fluid. In the best patterns the oscillations of the membrane are recorded by means of a beam of light which impinges on a mirror attached to the membrane and is reflected on to a moving photographic surface.

---

* FRANK. *Z. f. Biol.*, 1903, **44**, 445; 1908, **50**, 309; 1925, **82**, 49.

Manometers on these principles have been designed by Frank (Fig. 360), Wiggers, and Piper. Wiggers' manometer * is shown in Fig. 361. At its lower end cannulæ of various forms and sizes can be attached. At the middle of the tube is a short lateral tube to which, by a ball and socket joint B, is attached a tiny "segment capsule" A. Over the capsule is stretched a thick rubber membrane, to which on the straight side of the segment a small mirror is affixed. The entire capsule and its mirror can be adjusted in any direction by the handle D. A second small mirror E serves to project a base line, and is adjustable by the handle F. Alterations in the pressure within the manometer cause minute oscillations of the membrane, which can be recorded and magnified to any desired extent by means of a beam of light reflected from the mirror on to a moving photographic plate or paper (Fig. 362). The "figure of merit" of this manometer, *i.e.*, its own period of vibration, when filled with fluid, and with a thick membrane, is about 450 per second so that it can record with accuracy such rapid changes of pressure as occur even in the left ventricle.

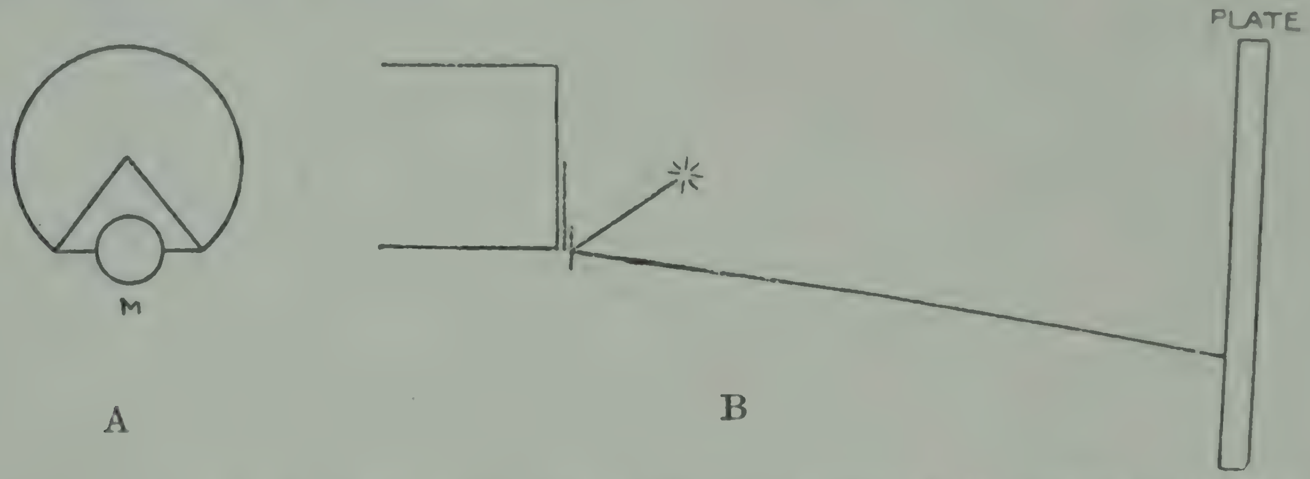


FIG 362. A, Magnified Diagram of a "Segment Recorder" covered with Membrane, to which is connected the minute Mirror M (front view).
B, Diagram showing how a Beam of Light is thrown on to the Mirror and then reflected on to a moving Photographic Plate.

In a further improvement a still more rigid membrane, in the form of a thin plate of beryllium copper, replaces the rubber one.†

INTRAVENTRICULAR PRESSURE CHANGES. In order to interpret these curves, we must have a simultaneous record of the pressures in the auricle and ventricle, or in the ventricle and aorta. Fig. 363 represents curves obtained from the

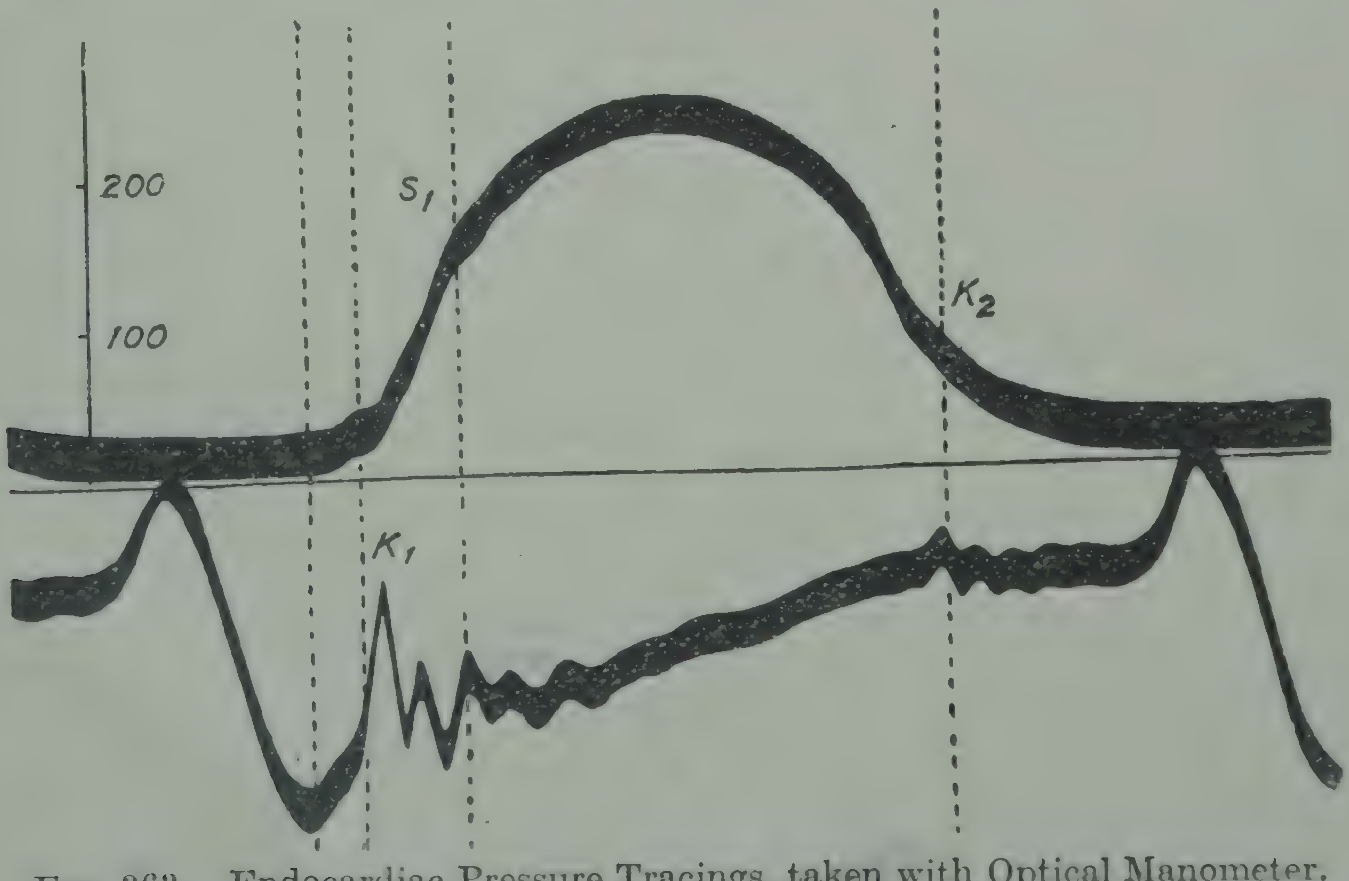


FIG. 363. Endocardiac Pressure Tracings, taken with Optical Manometer.
Simultaneous tracings, the upper from left ventricle and the lower from left auricle. To be read from left to right.

K_1 Closure of A.V. valves.
K_2 Opening of A.V. valves.
S_1 Opening of aortic valves.

---

* WIGGERS and BAKER. *J. Lab. Clin. Med.*, 1924, **10**, 54.

† HAMILTON *et al.* *Amer. J. Physiol.*, 1934, **107**, 427; GREGG *et al.* *Ibid.*, 1937, **118**, 399.

left ventricle and left auricle. On the up-stroke of the ventricular curve one or two elevations are usually visible, the most noticeable being the elevation marked S_1. This can be shown to correspond to the opening of the aortic valves.

The average course of the changes of pressure in the left heart during each beat is shown diagrammatically in Fig. 364. The cardiac cycle begins with the contraction of the auricle, which may cause a slight rise of pressure in the ventricle. As the auricular contraction dies away, the ventricular contraction begins at I. This causes a very rapid rise of pressure. Almost immediately after the beginning of the rise the auriculo-ventricular valves close, and the *isometric phase* begins. In this the pressure rises rapidly in the ventricular cavity without causing any change in its contents. Directly the pressure exceeds that in the aorta, the aortic valves open at the point marked II, where the *ejection phase* begins, and the aortic pressure now, from II to V, varies directly with the ventricular pressure. During the whole of this phase, the aortic pressure remains somewhat below the ventricular pressure, showing that blood is flowing continuously from ventricle into aorta. The ejection of blood is at first rapid, and the pressure in the ventricles continues to rise. As the heart empties, the amount of blood ejected into the aorta becomes less than that flowing away in unit time through the peripheral branches, so that the pressure begins to fall in the aorta and ventricle, even though ejection of blood is still going on. The ejection period may therefore be divided into two phases, that of *maximum ejection*, and that of *reduced ejection*. Usually this portion

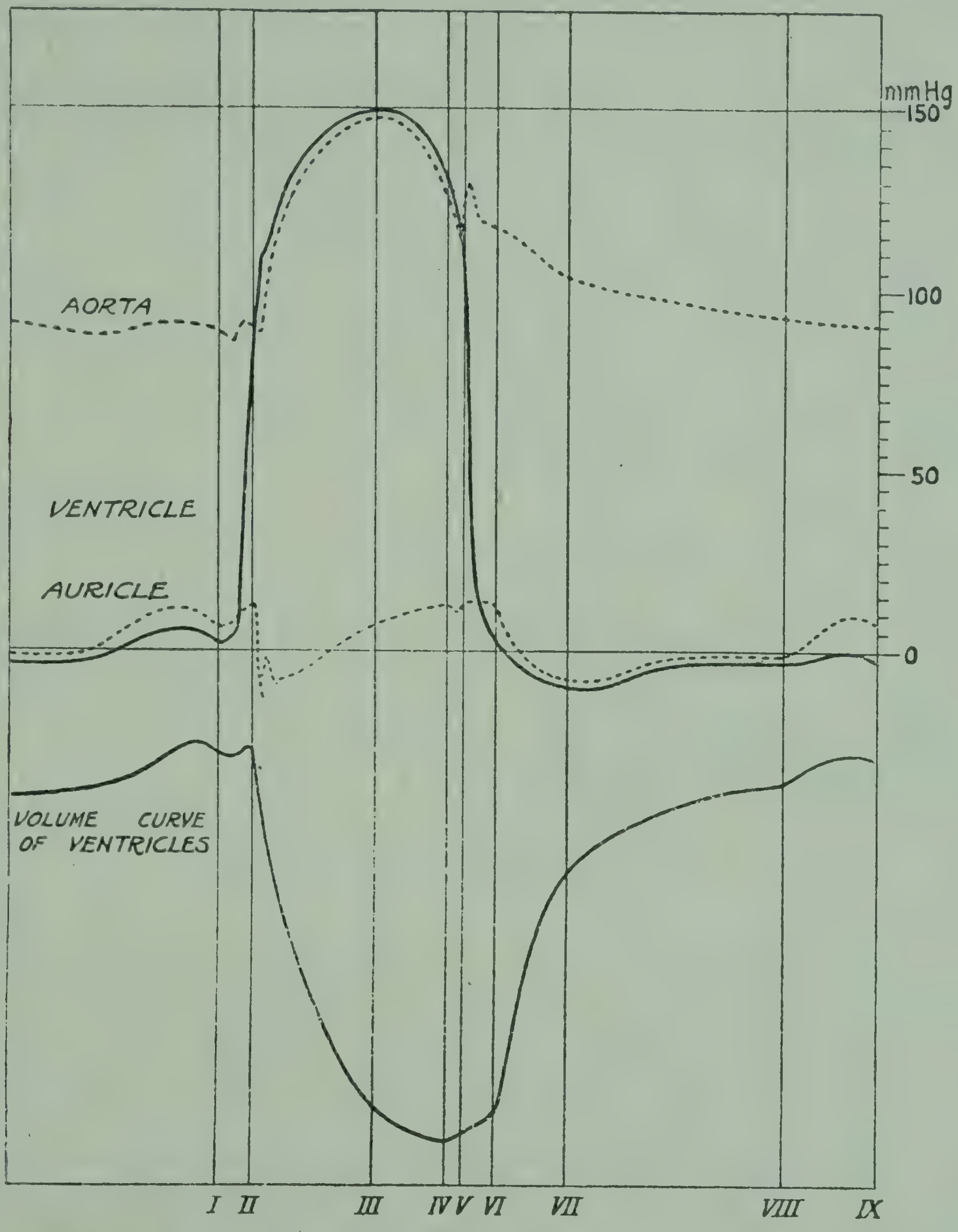


FIG. 364. Pressure Curves from Aorta, left Ventricle and left Auricle, together with Volume Curve of the two Ventricles during one complete Cardiac Cycle. (Modified from WIGGERS.)

of the curve forms a smooth arch, but under some conditions it may be flattened to form a plateau, of horizontal, ascending or descending slope. The ventricular muscle suddenly relaxes at the point marked IV, causing a sudden fall of pressure in the ventricle, and a slight fall in the aorta. This latter, however, is arrested almost at once by the closure of the aortic valves, marked by the sharp, depression, the *incisura* (V), in the aortic tracing. Between V and VI the relaxation is isometric, since all the valves guarding the orifices of the ventricle are closed. The pressure in the ventricle continues to fall, until at the point VI it drops below that in the auricle, and then the auriculo-ventricular valves open, allowing the inflow of blood from pulmonary veins and auricle. The pressure in the ventricles soon reaches zero and remains at or near this line during the greater part of diastole. With a big inflow there may be a slight rise towards the end of diastole, and a further rise on auricular contraction. If the chest is opened, the pressure in the ventricle never sinks below zero during any part of diastole. Naturally, under normal conditions, the pressure in the closed chest outside the heart is " negative " owing to the elastic retraction of the lungs, and may vary from — 3 to —30 mm. Hg, according as it is measured during normal expiration or during forced inspiration. This negative pressure may be transmitted to the interior of the heart ; hence, in the closed chest the pressure in the heart cavities during diastole may, as in Fig. 364, be negative, but never falls below the intra-pleural pressure which obtains outside them, *i.e.*, the heart does not actively exercise suction during diastole.

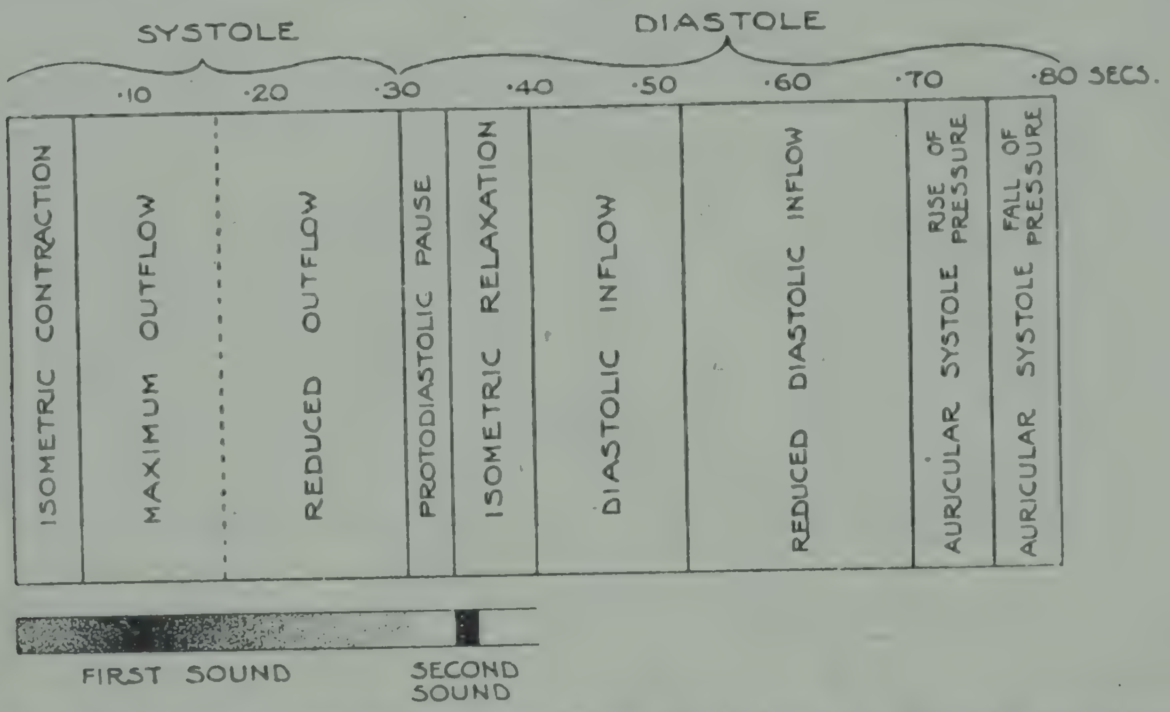


FIG. 365. Diagram showing average Time Relations of Events constituting Cardiac Cycle.

The time relations of these events naturally varies with the frequency of contraction of the heart. As a round figure, in a heart, beating 75 times p. m., a cardiac cycle will last 0·8 seconds. Based on this, the following may be taken as the average duration of the different phases in the left ventrclie (Wiggers) :—

(1) Last phase of ventricular diastole. A small rise of pressure due to contraction of the auricles, lasting 0·05 second. The whole auricular systole (rise and fall of intra-auricular pressure) takes about 0·11 second.

(2) The isometric phase, initiating systole, lasts 0·04 to 0·06 second. It varies but little with alteration in heart rate.

(3) The total duration of ventricular systole is about 0·3 second, varying from 0·33 to 0·23, as the heart rate quickens from 60 to 120 per minute.

(4) The diastolic phase may be divided into the following periods :—

(*a*) Protodiastolic phase (from beginning of relaxation to closure of aortic valves), 0·038 second.
(*b*) Isometric relaxation, 0·076 second.
(*c*) Rapid inflow phase, 0·113 second. } Joint diastole = 0·39 sec.
(*d*) Slow inflow phase, 0·163 second.

The relative durations of these last two phases will naturally vary according to the rate of the heart and the inflow of blood into the heart.

The average relation of these various phases of the heart cycle are represented diagrammatically in Fig. 365.

Similar but smaller pressure changes occur in the right side of the heart.

Intra-auricular Pressure Changes. Owing to the absence of valves between the right auricle and the venæ cavæ, changes of pressure within this cavity are transmitted along the veins. Observation shows that the jugular vein pulsates at least twice for each heart beat. The exact form of the pressure tracing in the auricles varies considerably, according to the inflow of blood and the state of filling of their cavities. A typical tracing with a moderate inflow of blood is given in Fig. 363, p. 558, and in this figure the relations of the different elevations in the auricular tracing to the intraventricular events can be made out ; in Fig. 366, the essential features of the curves are similar, and present the following features :

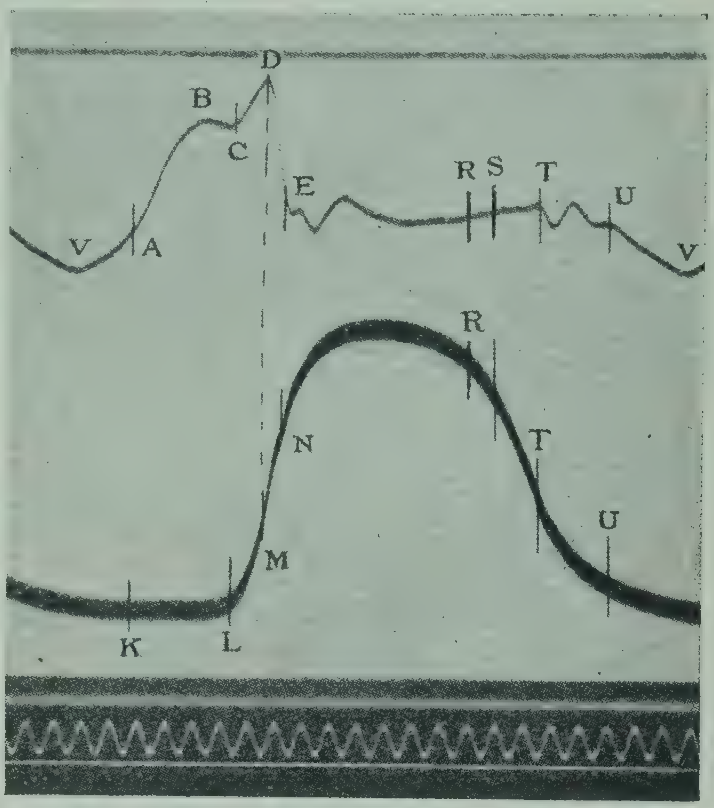


Fig. 366. Synchronous Left Auricular (upper) and Left Ventricular (lower) Pressure curves. (Wiggers.)

A, B, C = Auricular systole.
C, D = Second positive wave.
R = End of ventricular systole.
U, V = Rapid inflow phase.
Time = 0·02 second.

(1) The first positive wave, B (pre-systolic wave) corresponding to the auricular systole.

(2) The second positive wave D (first systolic wave) occupying the beginning of the ventricular systole. This is caused by the sharp closure of the A.V. valve.

(3) A negative wave of variable extent E largely due to the contracting ventricle pulling down the A.V. ring towards the apex, and so opening out the auricle " like a concertina."

(4) A third positive wave (second systolic wave) which may present secondary undulations. This rise of pressure is due to the gradual filling of the auricles while the auriculo-ventricular valves are still shut.

(5) A negative wave, not always well marked. At this point (U.V., Fig. 366) the ventricle is entirely relaxed and the auriculo-ventricular valves open, so allowing the blood to flow freely from the auricle into the ventricle.

The pressure rises in the left auricle somewhat higher than in the right auricle. In the latter the big veins act as a supplementary reservoir to the auricle, so that in no period of the cardiac cycle need the pressure in the latter chamber rise to any extent. In auricular tracings the heart sounds are often apparent as small oscillations in the curve (Fig. 363).

Changes in Form of the Heart. Under normal circumstances the heart in the unopened chest lies in the pericardium, and is also supported laterally by the lungs, which have very little influence on its shape. As the heart walls are flaccid during diastole, its shape will depend upon the position in which the heart is lying. Thus if the chest and the pericardium be opened and the animal be in the supine position, the heart during diastole will be flattened from before backwards. In this position, therefore, systole will be accompanied by a shortening in the lateral and vertical directions and a lengthening in the sagittal direction. During systole the heart, whatever its previous condition, takes the form of a truncated cone.

When the heart is freed from the pericardium, the obliquity of its fibres causes the apex to move forwards and to the right during systole ; this movement is normally prevented by the attachment of the pericardium to the central tendon of the diaphragm, so that the most movable part of the heart is then the base, which moves downwards towards the apex. This movement is determined partly by the shortening of the fibres of the ventricular wall partly by the lengthening of the great arteries as blood is forced into them from the ventricles. It was observed by Leonardo da Vinci that if three needles be passed through the chest wall so that their points lie, one in the base, one about the middle of the ventricles, and one in the apex of the ventricles, each ventricular systole is accompanied by a downward movement of the needle in the base of the heart, a slighter movement of the needle in the middle of the ventricles, and practically no movement at all of the needle which is thrust into the apex.

The Cardiac Impulse. Each ventricular systole causes a throb to be communicated to the chest wall, over a limited area of which it may be felt and seen, except in fat individuals. The region where the pulsation is most marked usually lies in the fifth intercostal space, a little to the median side of the left nipple. This position may shift with alterations in the work of the heart, as after exercise. The pulsation is spoken of as the *cardiac impulse* or "apex beat," though the apex of the heart is situated lower down. During diastole the ventricles form a flabby flattened cone lying against the chest wall and slightly deformed by the latter, especially when the subject leans forwards, or assumes the prone position. In systole the ventricles become rigid, assuming the form of a rounded cone. This sudden recovery of shape and hardening of the ventricular wall pushes out that part of the chest wall in immediate proximity to the ventricles, and so gives rise to the cardiac impulse. The cardiac impulse, by its characters of force, regularity, frequency and maintenance, gives valuable clinical information as to the occurrence and nature of the ventricular contraction.

The Heart Sounds. If we apply the ear, or listen with a stethoscope, to the front of a person's chest, we hear two distinct sounds accompanying each cardiac impulse, *i.e.* each systole, followed by a pause corresponding to the diastole. The sounds are often, though inadequately, compared to the syllables *lubb, dup*, the first sound being low-pitched and prolonged, the second sound higher and sharp. Thus the heart sounds may be represented: 1st, 2nd (pause), 1st, 2nd (pause), etc.

The *first sound* is best heard at the situation of the cardiac impulse. It is caused by the sudden closure of the auriculo-ventricular valves and the sudden increase in tension upon them and in the ventricular walls, whereby all these structures are caused to vibrate. It was formerly believed that the contraction of the ventricular muscle contributed a component to the first sound, but any such component is probably small, because little sound is produced if means are adopted to prevent alteration in tension of the ventricular walls and of the auriculo-ventricular valves.* The first sound is reduplicated if the two ventricles contract asynchronously (as in bundle branch block).†

The *second sound* is heard most loudly just over the second right costal cartilage, *i.e.* where the aorta lies nearest the surface. It comes at the end of ventricular systole, as determined by the cardiac impulse, and is synchronous with the closure of the aortic valves. It is, in fact, caused by the vibrations set up in the valves, and in the walls of the aorta near them, consequent upon their sudden shutting and stretching directly the heart ceases to contract and to force the blood into the aorta. If the valves be hooked back in an animal by means of a wire passed down a carotid artery, the second sound disappears and is replaced by a murmur caused by the blood rushing back into the ventricle at the end of the systole.

The pulmonary and aortic valves generally close simultaneously. In some cases, however, the aortic may close slightly before the pulmonary, giving rise to a " reduplicated second sound."

---

* Eckstein. *Amer. J. Physiol.*, 1937, **118,** 359; Lewis and Dock. *J. Amer. Med. Assoc.*, 1938, **110,** 271.

† Braun-Menéndez and Solari. *Arch. Int. Med.*, 1939, **63,** 830.

The distinctness with which the vibrations causing the heart sounds are heard, and also their character and duration, depends primarily on the physical conditions for conduction, such as the thickness of the chest wall, the properties of the stethoscope, etc., but other things being equal, the loudness of the first sound depends on the abruptness and force with which the ventricle contracts, while the loudness of the second sound varies with the force with which the aortic valve is closed. The slower vibrations of the sounds are suppressed by stethoscopes and also by the ear, and so tend to be lost.

THE GRAPHIC RECORD OF THE HEART SOUNDS. Various methods have been devised for graphically recording the heart sounds. They have been recorded directly by delicate membranes bearing mirrors. The usual method is to convey the heart sounds to a microphone, which is thus set into vibration. The alterations in current so produced are recorded by a string galvanometer, and a simultaneous record is usually made of the jugular pulse and E.C.G. A record of the heart sounds is shown in Fig. 367. It reveals vibrations which are usually inaudible. The first sound, as recorded, may last for as long as 0·165 of a second, though the audible part, using a stethoscope, may be less than half that duration. Its frequency ranges from 25 to 40 per second. Its early, slower vibrations (A) the "auricular sound," synchronise with auricular systole, and are normally inaudible. The readily audible vibrations correspond to the isometric and early ejection phases of ventricular systole.

The audible portions of the second sound correspond approximately to closure of the semilunar valves, at about the beginning of the V-wave of the venous pulse: the frequency is about 50 per second, and duration 0·08 to 0·14 second, according to conditions of recording.

THE THIRD HEART SOUND. A third heart sound is described * as occurring in certain individuals during diastole, shortly after the second sound. It may be mistaken for a reduplicated second sound. It is softer and of a lower pitch than the second sound, and is heard most distinctly over the apex beat. It is probably due to the vibrations set up in the fluid itself, or in the auriculo-ventricular valves, by the sudden inrush of blood from auricles to ventricles at the beginning of diastole.† The sound is shown in Fig. 367.

CARDIAC MURMURS. When a fluid escapes through a narrow orifice into a wider space, vibrations are set up in the fluid and may cause sound. The same sort of vibration may be set up in the large vessels or in the heart, whenever the blood passes rapidly through a narrow orifice into a wider space.

In the normal resting individual, sounds produced in this way, though detectable by delicate recorders, are so slight that they are inaudible; examples are seen in the third sound and in the auricular sound already mentioned; under abnormal conditions, as after diseases affecting the valvular orifices of the heart, or after exercise, they may be heard with ease. These murmurs, or *bruits* as they are called, are of paramount importance in enabling the medical man to form a judgment as to the condition of the different valves of the heart. Thus injury to an aortic valve, so as to allow of leakage during diastole, allows a small amount of fluid under high pressure to pass from the aorta into the relaxed ventricle. On listening to the chest of a man with such a lesion, this regurgitation during diastole is heard as a rushing sound occurring in the place of, or continuing, the second sound up to the beginning of the next first sound, which denotes the commencement of systole.

In many cases, the disease which occasions the incompetence of the aortic valve is followed by cicatrization, so that the orifices can never become thoroughly patent or properly closed. Under such circumstances vibrations will also be set up in the current of blood as it escapes through the narrow orifice into the aorta during systole; and on listening to the chest over the second right costal cartilage, a " to and fro " bruit is heard composed of a systolic immediately followed by a diastolic murmur. In the same way, incompetence of the mitral valve, or dilatation of the mitral orifice, gives rise to a murmur which lasts during the whole of the ventricular contraction and is therefore

---

* EINTHOVEN. *Pflüger's Arch.*, 1907, **120**, 31.

† GIBSON. *Lancet*, 1907, ii, 1380.

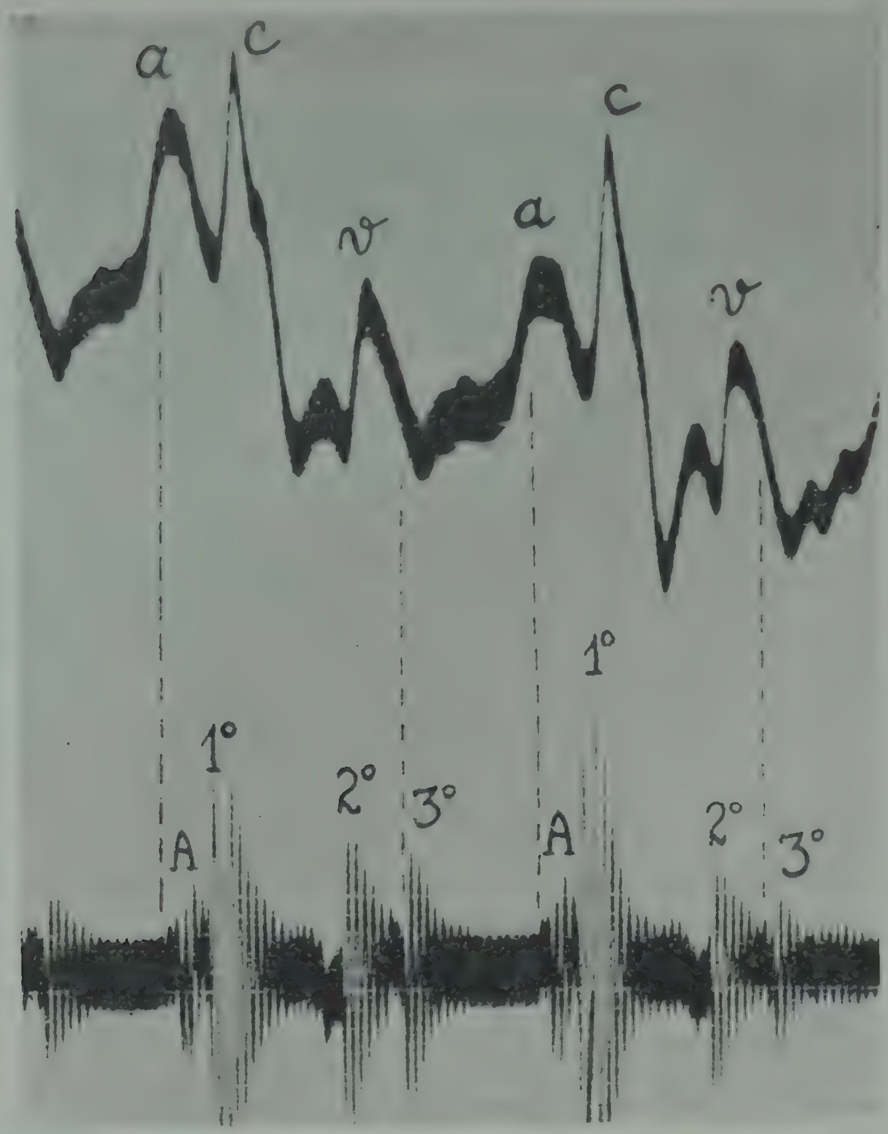


FIG. 367. Jugular vein pulse and heart sounds in man. In addition to the audible first and second sounds, the normally inaudible auricular sound (A) and the third sound are shown. (E. BRAUN-MENÉNDEZ and O. ORÍAS. *Rev. Argent. Cardiol.*, 1934, **1**, 101.)

systolic in relation. Such a murmur is heard best over the area of the cardiac impulse, and is also transmitted backwards, so that it can be heard on listening at the back of the patient. A narrowing of the mitral orifice in consequence of cicatrization of the valves will set up a resistance to the flow of blood from left auricle to left ventricle. The auricle becomes hypertrophied, its contraction prolonged, and the passage of blood through the contracted orifice gives rise to an auricular murmur which is heard on listening over the apex beat as a *presystolic bruit*. This bruit is easily distinguished from a systolic murmur by noticing that it runs up to and ends with the cardiac impulse, whereas a systolic murmur does not begin until the cardiac impulse commences.

For Reference

BRAUN-MENÉNDEZ. The Heart Sounds in Normal and Pathological Conditions. *Lancet*, 1928, ii, 761.

ORÍAS and BRAUN-MENÉNDEZ (1939). The Heart Sounds in Normal and Pathological Conditions. (Oxford Med. Publ.).

The Human Electrocardiogram. When an excitable tissue is immersed in saline solution, the electrical changes led off during excitation are not of the simple diphasic or monophasic nature obtained when the tissue is in air.* Similarly, if we lead off any two parts of the heart's surface to an oscillograph, we record potential changes which are caused partly by events occurring in the muscle just underlying the electrode, partly by changes occurring at a distance and transmitted by the intervening muscle and blood acting simply as a moist conductor.† These two kinds of effect may be alluded to as direct and indirect. If we lead off, not from the heart itself, but from neighbouring tissues in contact with it, we shall still obtain the indirect effects of the electrical changes at each beat; and so in like manner these can be obtained when very various couples of points on the surface of the intact animal are led off to the galvanometer. In man, where the heart lies asymmetrically, with its electrical axis varying during systole, but roughly lying in an oblique direction pointing downwards towards the left, it is usual to start with the three classical "leads," viz. right arm—left arm (lead I); right arm—left leg (II); left arm—left leg (III). More recently, various other leads have been added to elicit further information, *e.g.* chest leads (IV) to right arm, left arm, or left foot (IV.R; IV.L; IV.F), or various precordial leads (C) (*e.g.* CR; CL; CF), or a CV lead ("unipolar lead"), the V lead being obtained by connecting R, L and F together as one lead.*

---

* CRAIB, *Heart*, 1927, **14**, 71; *J. Physiol.*, 1928, **66**, 49; GILSON and BISHOP, *Amer. J. Physiol.*, 1937, **118**, 743; WILSON, MACLEOD and BARKER, *J. Gen. Physiol.*, 1933, **16**, 423.

† LEWIS, MEAKINS and WHITE. *Phil. Trans. Roy. Soc.*, 1914, **205B**, 375.

In all cases, each pair of leads, from suitable electrodes, are led to a string galvanometer or cathode ray oscillograph (Fig. 45, p. 118). In the record, or *electrocardiogram*, the time scale is uniform, and the electric potential scale always the same (1 cm./mv.): negativity of F or of C, is in a downward direction, and of R or L upward, *i.e.* excitation of the base of the heart, roughly speaking, gives an upward movement, and that of the apex a downward one. A typical electrocardiogram (E.C.G.) is shown in Fig. 368. The different points in the tracing are designated by the letters P, Q, R, S, T, first applied by Einthoven and retained because they do not commit us to theoretical interpretation of the curves, which, in view of the reasons given above, would be difficult; P, R and T are usually upward waves, Q and S, downwards. Such interpretation as we have is the result of direct experimental observation, which shows that the electrocardiogram is a composite picture of different events in the heart; that P is certainly due to the auricular excitation, and Q marks the beginning of the ventricular excitation. The P wave averages 0·08 sec. and precedes auricular contraction by 0·02 sec. Q, R, S varies normally, with heart rate, from 0·04–0·08 sec., and represents the spread of excitation in the ventricles. It precedes their contraction. Q—T (0·43 sec.) gives duration of systole. The A—V. interval is given by the distance between P and R

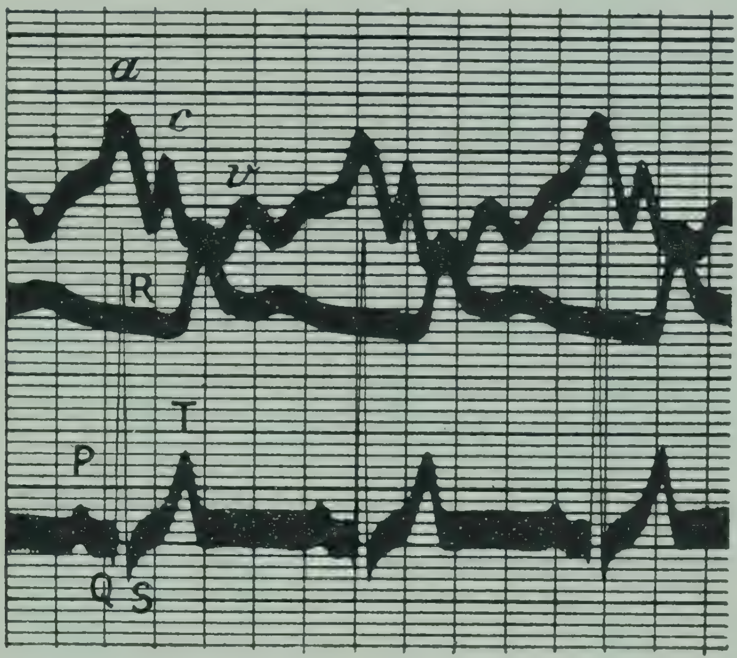


FIG. 368. Jugular venous pulse, radial arterial pulse and electrocardiogram in man. Vertical lines = 0·2 second: horizontal lines 0·1 mv. on E.C.G. The jugular waves *a*, *c*, *v* are marked; also the E.C.G. waves P, Q, R, S, T. Note that, due to delay in pulse transmission, the P precedes *a*, and *c* precedes the radial upstroke. The interval between auricular and ventricular systole, as measured by *a*—*c*, is slightly greater than that obtained from P—R (0·15 secs.). The summit of R is about 0·03 sec. before ventricular systole begins; the summit of T is at height of systole and end of T about 0·03 sec. before second sound. (LEWIS. *Diseases of the Heart*, 1933. Macmillan.)

---

* NAHUM and HOFF. *Amer. J. Physiol.*, 1948, **155** 215.

(0·12–0·2 sec.), the total duration of the excitatory condition in the ventricle by the distance between Q and T. The highest wave, R, ranges normally from 0·7 to 1·8 mv. The T wave (0·27 sec.) ends at the termination of ventricular systole.

Quite apart from the objections outlined above, we should not expect, in view of the mechanism of the propagation of the excitatory wave in the ventricle, that the electrogram obtained in this indirect fashion would be easy of interpretation, at any rate so far as regards the course of the wave through the ventricular muscle. It must be remembered that the ventricular muscle is excited almost simultaneously at all points, in consequence of the rapid propagation of the wave of excitation along the branches of the A.V. bundle. The exact interval at which the excitatory process arrives at different points of the surface of the dog's ventricle after the upstroke of R in the whole electrocardiogram is shown, for example, in Fig. 369.

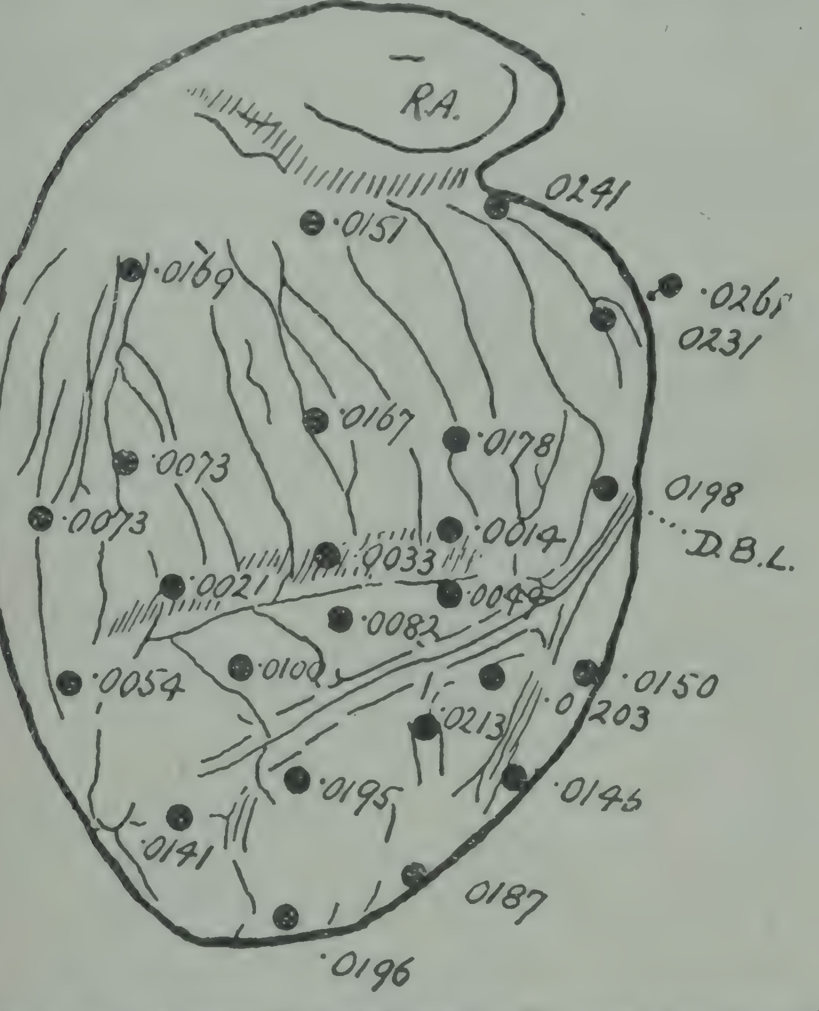


FIG. 369. Diagram showing the Time of Arrival of the Excitatory Wave at different points on the surface of the Ventricles in the Dog. (LEWIS and ROTHSCHILD, *Phil. Trans. Roy. Soc.*, 1915, **206B**, 181.)

The exact form and time relations of the ventricular part of the electrocardiogram may be altered in disease, or by interruption of one or other of the main branches of the A.V. bundle, because then the order of spread of the wave of excitation becomes different.

DEXTRO- AND LÆVOCARDIOGRAMS. When the right branch of the A.V. bundle is blocked or cut, the excitatory wave will reach the left ventricle first, and if the left branch is divided, then the right ventricle will contract first. In either case the normal composite electrocardiogram will be departed from, and the resulting record is called lævocardiogram (*i.e.* right branch blocked), or dextrocardiogram (left branch blocked) respectively, because it shows respectively the effect of an initial and preponderating contraction of left or right ventricles. A similar record will be obtained if a premature contraction is provoked by a stimulus applied to the appropriate ventricle. The results of direct stimulation of the ventricles in man * show that the two curves have the following characters :—

Dextrocardiogram. *Lead I*, initial deflection (R) upwards ; *Lead III*, initial deflection (R) downwards.

Lævocardiogram. *Lead I*, initial deflection (S) downwards ; *Lead III*, initial deflection (R) upwards.

(Owing to incorrect inferences from other species, the earlier investigators believed that the deflections were the opposite of those stated above.)

The electrocardiogram, moreover, is of considerable use clinically for the determination of the relation between the auricular and the ventricular contractions. Thus, in heart block there will be P waves independently of Q, R, S waves ; in a partial block there will be delayed conduction time, as shown by increased P—R interval (over 0·2 sec.) ; in auricular fibrillation the P waves are replaced by small rapid waves, etc.

For Reference

LEWIS (1925). The Mechanism and Graphic Registration of the Heart Beat. (London : Shaw & Sons).

KATZ (1928). The Significance of the T wave in the Electrogram and Electrocardiogram. *Physiol. Rev.*, **8**, 447.

KATZ (1947). The Genesis of the Electrocardiogram. *Physiol. Rev.*, **27**, 398.

---

* BARKER, MACLEOD and ALEXANDER. *Amer. Heart J.*, 1930, **5**, 720.

Filling of the Heart in Diastole. The heart during diastole exerts no active suction on the blood in the veins. Its filling depends on the existence of a pressure within the veins greater than that in the auricles and ventricles * ; the greater the pressure within the large veins the more rapidly will the blood enter the heart during diastole. In this process an important part is played by the mechanical conditions existing in the thoracic cavity. Owing to the elasticity of the lungs and to the fact that they are constantly tending to contract, the pressure in the thorax is less than that of the external atmosphere. At the end of expiration this difference amounts to about 5 mm. Hg, rising to 9 mm. Hg at the end of inspiration and to 30 mm. Hg at the end of a forced inspiration. This " negative pressure " chiefly affects the thin-walled veins and auricles, and the right ventricle when it is relaxed. On the other hand, the veins outside the thorax are exposed to a pressure which is a little above that of the atmosphere. The respiratory movements, by causing an alternating suction on the walls of the great veins and heart, act like an accessory pump and cause an aspiration of blood into the heart with each inspiration.

If the pressure within the thorax be raised enough to cause a sufficient positive pressure on the big veins and auricles, the return flow of blood into the heart is suspended. Thus during extreme muscular effort with the glottis fixed a positive pressure is produced in the thorax. The impaired circulation is rendered evident by the engorgement of the superficial veins. Weber showed that, by a forcible expiration with the glottis closed, the pulse might disappear at the wrist and the circulation be brought to a standstill for such a time that loss of consciousness might supervene.

The whole heart is enclosed in the tough, fibrous, inextensible sac formed by the pericardium. Under normal circumstances the heart, even during diastole, does not entirely fill the sac, increased rate of filling being compensated by a reflex quickening of the heart rate. Under extreme stress, when both inflow and resistance to outflow are increased, or when the heart muscle is in a bad condition from fatigue or disease, the heart may expand during diastole so as entirely to fill the pericardium, which then prevents any further distension, and, by limiting inflow, limits also the stroke volume of the heart. Thus, the size of the pericardium determines the maximum possible output per beat. If the pericardium be removed, the heart may distend in diastole still further, the only limit to its filling being set by the resistance of the heart wall. Under these conditions the contraction may not empty the ventricle and the heart may fail.

The normal filling of the heart during diastole can be prevented by anything which hinders its expansion, such as the presence of fluid in the pericardial cavity. This happens in puncture of the left ventricle of the heart, as by rapier wounds. The same effect may be produced if oil be injected into the pericardium in a dog with the chest open ; when the pressure of the oil in the pericardial cavity is equal to that in the veins to the heart, the arterial blood pressure falls to zero and death ensues.†

It is under such conditions that the *pulsus paradoxus* may be observed, in which the pulse becomes small or absent during inspiration. This is due to the fact that, when the pericardial sac is filled with fluid under a positive pressure, the auricles and ventricles are much less affected by respiratory variations of intrathoracic pressure. The big veins, however, are still subject to these variations, so that during inspiration the blood tends to accumulate within them and not pass on into the heart. This organ therefore has less blood to discharge, and there is a diminution or absence of the pulse during inspiration. (Katz and Gauchat.)

Systolic Output of the Heart. The rate at which the blood circulates round the body depends on the quantity of blood which leaves the left heart per unit of time. In normal fasting men at rest the output is about 5 to 6 l. per minute, or 2·8 ± 0·3 l.per minute per sq. metre of body surface, this latter figure being called the *cardiac index*. It falls slightly with advancing age, from forty to fifty-five, and then more slowly.* It is not surprising that the output depends, among other things, on the body surface, being, in fact, for most mammals at rest, usually 2 to 3 l./sq.m./min. The output of the right ventricle must over any length of time be equal to that from the left ventricle, otherwise the blood would accumulate on one or other side of the heart. It must be remembered, however, that the coronary arteries, supplying the wall of the heart itself, leave the aorta at its very commencement, just beyond the aortic valves. The quantity of blood passing per unit of time along the aortic arch is therefore less than that leaving the left ventricle by the amount of blood which traverses the coronary vessels. The true output of each ventricle is therefore given, not by the flow through the aorta, but by the flow through the pulmonary artery.

---

* HENDERSON. *Amer. J. Physiol.*, 1906, **16**, 325.

† KUNO. *J. Physiol.*, 1917, **51**, 221.

Methods for experimentally determining the cardiac output fall into two classes. In the first the *stroke volume*, *i.e.* the volume changes in the ventricles at each beat of the heart is determined. A second method consists in the determination of the total volume of blood leaving the right or left ventricle in the course of a given time, say one minute—*the minute volume*. If this amount be divided by the number of heart beats in the same time, the stroke volume is obtained.

METHODS OF DETERMINING OUTPUT

(i.) **Cardiometric Method.** For recording the stroke volume in animals with opened thorax, a cardiometer † of the shape shown in Fig. 370, is slipped over the heart. To the margin of the opening is tied a rubber diaphragm with a hole in it, which accurately fits the heart and lies in the auriculo-ventricular groove. The tube of the cardiometer is connected with a piston recorder. During diastole the ventricles increase and during systole they diminish in volume. The change in volume at each beat gives the combined output of right and left ventricles, and must be divided by two in order to give the output of either ventricle separately. A cardiometer record is shown in Fig. 376. The disadvantages of this method are the open chest, and that the inertia and friction of the moving parts deform the curve.

(ii.) **Application of the Fick Principle.** If we know the difference between the oxygen content of the right heart and arterial bloods, a determination of the oxygen intake by the lungs over a given interval of time will tell us how much blood has passed through the lungs during that time. We thus get the *minute volume*. The mixed venous blood is drawn from the right ventricle. Thus ‡ in one case, in a horse weighing 360 kilos, 2,733 c.c. of oxygen were taken up in the lungs per minute, while the arterial blood contained 10·33 per cent. more oxygen than the venous blood. Since every 100 c.c. of blood that passed through the lungs had taken up 10·33 c.c. of oxygen, and 2,733 c.c. had been taken up in the course of a minute, it follows that

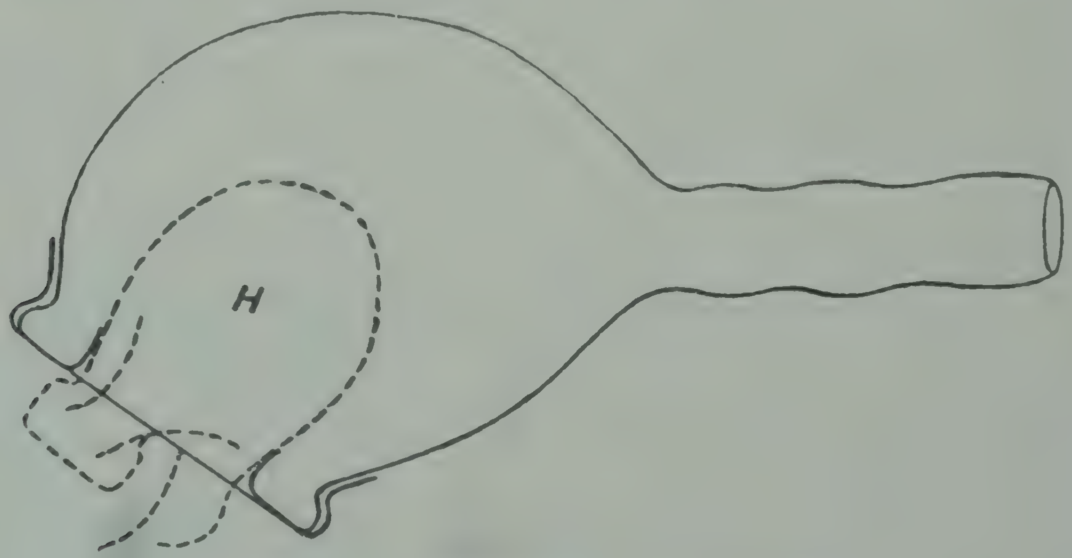


FIG. 370. Henderson's Cardiometer.

$$\frac{100 \times 2{,}733}{10{\cdot}33} = 26{,}457 \text{ c.c.}$$

of blood must have passed through the lungs in the time. Instead of oxygen differences, CO_2 differences may be taken, and the total output of CO_2 by the lungs determined.

---

* LEWIS. *Amer. J. Physiol.*, 1938, **121**, 517.

† HENDERSON. *Amer. J. Physiol.*, 1906, **16**, 325.

‡ ZUNTZ and HAGEMANN. "Stoffwechsel des Pferdes," 1898, p. 372. (Berlin, Parey.)

DETERMINATION OF CARDIAC OUTPUT IN MAN. The direct Fick principle can be applied to man. It involves the passing of a catheter into the right auricle by way of an appropriate vein, usually in the bend of the elbow,* and is not difficult (Fig. 371).

(iii.) **The Indirect Fick Principle.** It is also possible, by indirect methods, to find the CO_2, or O_2, in the arterial and venous bloods without the use of operative procedures. In Field's method, the tension of CO_2 in the arterial blood is obtained by determining the amount of this gas in ordinary alveolar air, since the alveolar and arterial tensions of CO_2 are identical. This tension must be converted into percentage content by reference to a CO_2-dissociation curve of the same blood (Ch. XL).

The content of CO_2 in the venous blood coming into the lungs is determined by a procedure which brings the alveolar air up to the same tension. The subject having expired deeply, takes, from a 5 l. bag, three deep breaths in succession of a mixture containing about 94 per cent. oxygen and 6 per cent. CO_2, *i.e.* just above what is known to be the tension of CO_2 in arterial blood. The large amount of oxygen is supplied in order to prevent errors due to anoxæmia. These three breaths are taken within fifteen seconds, and the last one is held for a few seconds. As the total circulation time in a man at rest is assumed to be about thirty seconds, there is no time for the blood which has taken up CO_2 from the lungs to come back to the lungs, *i.e.* equilibrium is established between the air in the alveoli and the continually renewed venous blood arriving at the alveoli by the pulmonary artery. After holding the third breath the alveolar air is collected and analysed so that its tension of CO_2 can be calculated.† This is converted into percentage of CO_2 in the venous blood by reference to the CO_2-dissociation curve, as was done for arterial blood. In this way we find the difference between the CO_2 content of the blood leaving and of the blood entering the lungs. Before these determinations are carried out, the total CO_2 given off by the individual is measured by collecting his expired air. Suppose that the difference in CO_2 content of arterial and venous blood was 40 c.c. per litre blood, while the total amount of CO_2 given off from the lungs in the course of a minute was 200 c.c. the total amount of blood would be $\frac{200}{40} = 5$ litres, which would thus be the "minute" volume of the heart. In other methods the arterio-venous oxygen difference is determined.

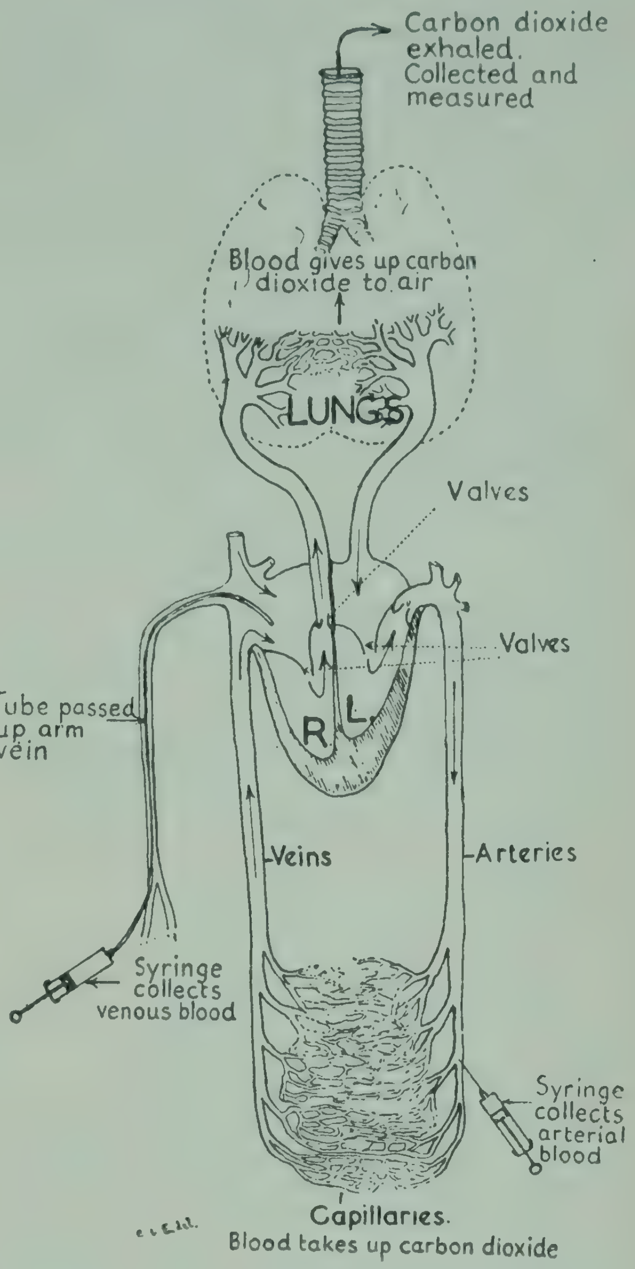

Carbon dioxide exhaled. Collected and measured
Blood gives up carbon dioxide to air
LUNGS
Valves
Valves
Tube passed up arm vein
R
L
Veins
Arteries
Syringe collects venous blood
Syringe collects arterial blood
Capillaries.
Blood takes up carbon dioxide

FIG. 371. Diagram to illustrate application of the Fick principle to man. Either CO_2 differences, as in figure, or O_2 differences can be utilised: in practice the latter is preferred. (By courtesy of Messrs. SAMPSON LOW, MARSTON & Co.)

Krogh and Lindhard ‡ devised a method for determining the blood flow through the lungs by finding how much nitrous oxide is taken from a mixture of nitrous oxide and air with which the lungs are filled. Knowing the solubility coefficient of nitrous oxide in blood at the tension and temperature at which this gas is present in the pulmonary alveoli, and determining the amount of this gas which has disappeared from the mixture breathed during a certain time, it is possible to estimate the volume of blood which has passed through the lungs during the same time. In another method the soluble gas used is acetylene.§

(iv.) **By Calculation.** It has been claimed that the cardiac output can be computed from the arterial blood pressure and the pulse wave velocity.||

(v.) **The Ballistocardiogram.** When the heart expels its blood into the arteries, a recoil (comparable with that of a rifle, together, sometimes, with the rifleman, when a bullet is fired) is given to the heart and to the whole body. If one stands on a delicate spring

---

* COURNAND and RANGES. *Proc. Soc. Exp. Biol. Med.*, 1941, **46**, 462; MCMICHAEL and SHARPEY-SCHAFER. *Brit. Heart J.*, 1944, **6**, 33.

† There are differences of opinion about the justifiability of the assumption that equilibrium is reached, cf. ISRÄELS and LAMB, *J. Physiol.*, 1929, **67**, 315; HAMILTON *et al.*, *J. Physiol.*, 1930, **70**, 244.

‡ KROGH and LINDHARD. *Skand. Arch. f. Physiol.*, 1912, **27**, 100; KROGH, *ibid.*, p. 227; LINDHARD, *J. Physiol.*, 1923, **57**, 17.

§ MARSHALL and GROLLMAN. *Amer. J. Physiol.*, 1928, **86**, 117; GROLLMAN, *ibid.*, 1929, **88**, 432.

|| BAZETT *et al.* *Amer. J. Physiol.*, 1935, **113**, 312.

weighing-machine, the pulsation due to this ballistic recoil, is clearly shown as a pulsation. A curve, the ballistocardiogram, can be recorded by having the subject lie on a suspended support movable in the long axis of the body.* It shows important changes with altered heart output.

CARDIO-PNEUMATIC MOVEMENTS. Since the ventricle during systole becomes smaller, the space thus provided in the chest cavity is taken up by an expansion of the veins, auricles, and lungs. To this systolic diminution of intrathoracic pressure are due the "cardio-pneumatic" movements. These are recorded by connecting one nostril to a tambour, while the other nostril and the mouth are kept closed. If a carotid pulse tracing be taken at the same time, it will be found that there is a fall of the pressure in the nasal cavity, synchronous with the rise of the pressure in the arteries, both being due to the expulsion of blood from the heart.

For Reference

CHRISTIANSEN (1937). Das Herzminutenvolumen. *Ergebn. d. Physiol.*, **39**, 348.
GROLLMAN (1932). Cardiac Output. (Baillière, Tindall and Cox.)
HENDERSON (1923). Volume Changes of the Heart. *Physiol. Rev.*, **3**, 165.
MARSHALL (1930). The Cardiac Output of Man. *Medicine*, **9**, 176.
STARR (1946). The Ballistocardiograph. *Harvey Lects.*, 1946–7, 194.

---

* STARR and RAWSON. *Amer. J. Physiol.*, 1938, **123**, 195.

CHAPTER XXXII

FACTORS WHICH MODIFY THE HEART'S ACTION

The performance of the heart is subject to considerable alterations, due to the operation of numerous factors capable of affecting either the frequency or the extent of its contractions. With the heart *in situ* these factors are difficult to analyse, owing to their number and complex interactions, particularly those of the mutual effects of the systemic circulation and the heart beat upon one another. In order to study the conditions which can control the filling and emptying, and also the rate, of the heart in the absence of the complications normally introduced by the presence of a delicately poised and constantly changing peripheral circulation, it is of importance to be able to replace the latter by an artificial circulation, the resistance of which can be controlled at will. This can be effected by the use of the heart-lung preparation.

The Heart-Lung Preparation. With the use of this preparation, devised by Starling,* it is possible to study the action of the cat or dog heart under very varying conditions of systemic arterial resistance, venous pressure, rate of filling, or temperature. One arrangement of the apparatus is shown in Fig. 372. Artificial respiration being maintained, the chest is opened under an anæsthetic. The arteries coming from the arch of the aorta—in the dog, the brachio-cephalic and the left subclavian—are then ligatured, thus cutting off the whole blood supply to the brain. Cannulæ are placed in the brachio-cephalic artery and the superior vena cava. The cannulæ and the large venous reservoir are filled beforehand with defibrinated blood. The descending aorta, azygos vein, and inferior vena cava are ligated. The only path left for the blood is by the ascending aorta and the cannula *CA* in the brachio-cephalic artery. The arterial cannula communicates with a

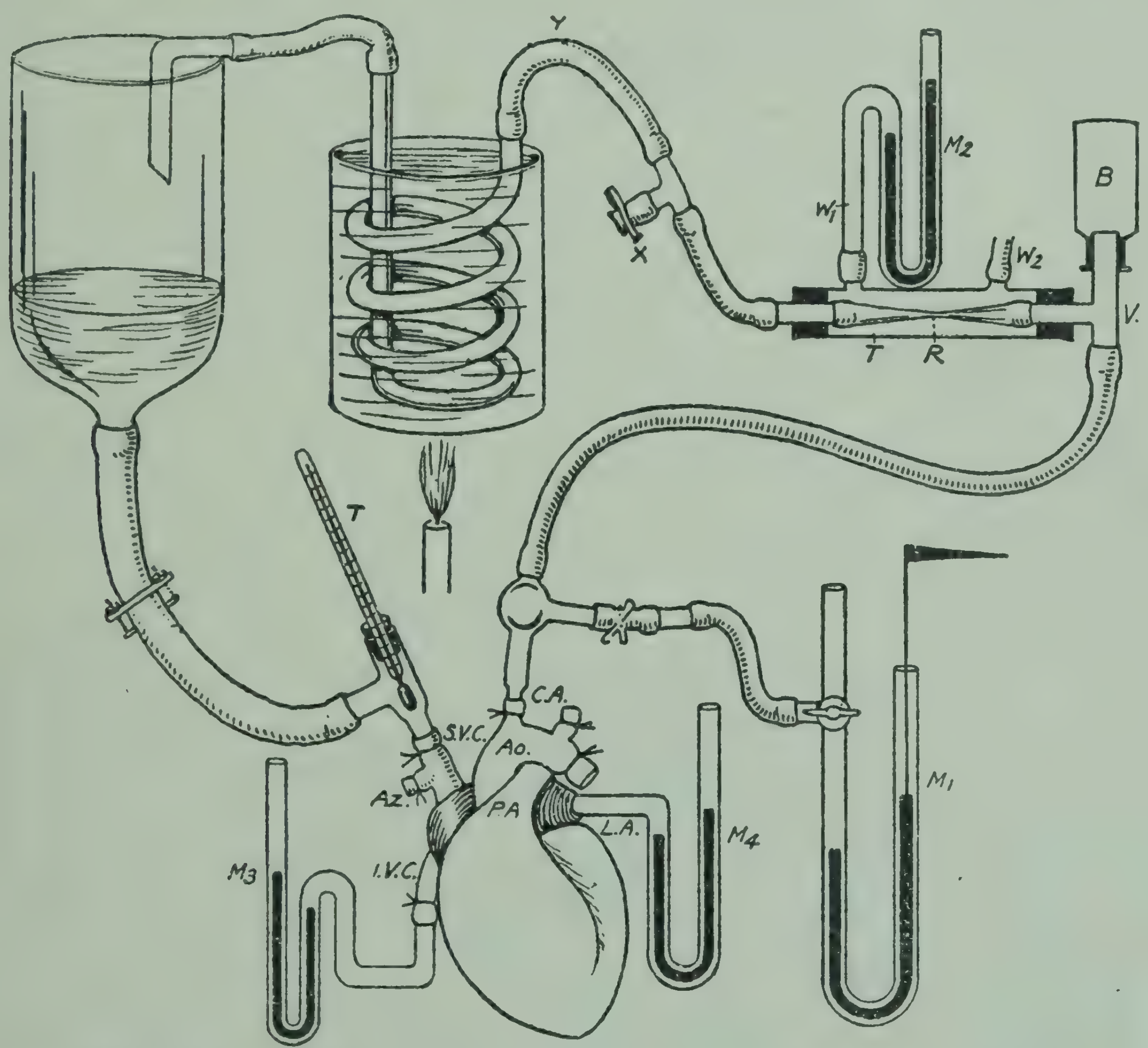


FIG. 372. Arrangement for the "Heart-lung Preparation."
The different parts are not drawn to scale, and the lungs are not shown. (STARLING.)

---

* JERUSALEM and STARLING. *J. Physiol.*, 1910, **40**, 285, modified by KNOWLTON and STARLING, *ibid.*, 1912, **44**, 206; and PATTERSON and STARLING, *ibid.*, 1914, **48**, 357. A closed circuit system has been devised by DALY, *J. Physiol.*, 1925, **60**, 103: an arrangement for the rat is given by CRUICKSHANK, *J. Physiol.*, 1945, **103**, 417.

mercurial manometer M^1 to record the mean arterial pressure, and passes to another T-tube *V*, one limb of which projects into a bottle *B*. The air in this vessel serves as an elastic resistance, taking the part of the resilient arterial wall. The other limb of the T-tube

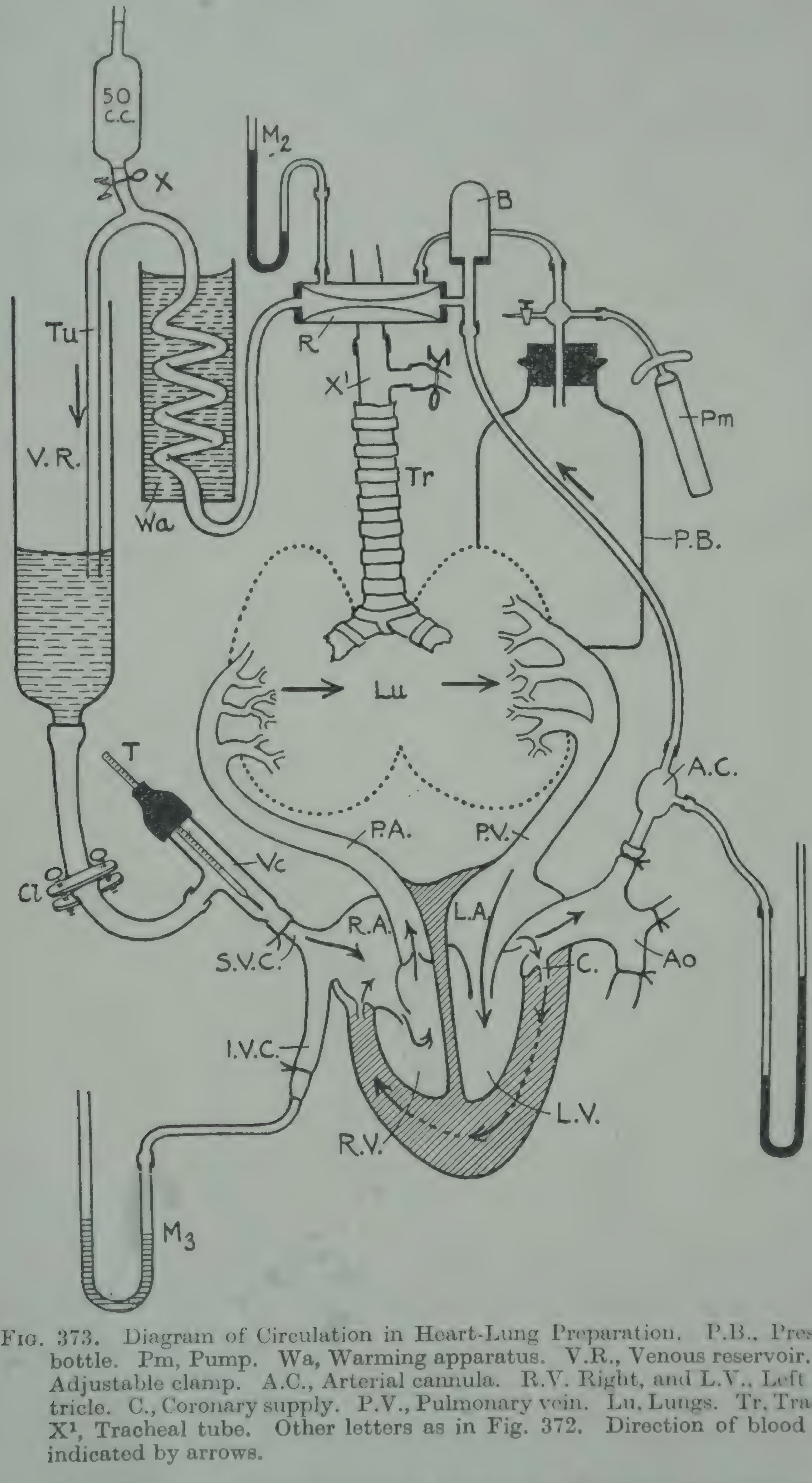


FIG. 373. Diagram of Circulation in Heart-Lung Preparation. P.B., Pressure bottle. Pm, Pump. Wa, Warming apparatus. V.R., Venous reservoir. Cl, Adjustable clamp. A.C., Arterial cannula. R.V. Right, and L.V., Left ventricle. C., Coronary supply. P.V., Pulmonary vein. Lu, Lungs. Tr, Trachea. X^1, Tracheal tube. Other letters as in Fig. 372. Direction of blood flow indicated by arrows.

passes to the resistance *R*. This consists of a thin-walled rubber tube (part of a rubber fingerstall) which passes through a wide glass tube provided with two lateral tubulures W_1, W_2. One of these, W_2, is connected with an air reservoir into which air can be pumped.

When air is injected into the outer tube, the tube *R* collapses, and will remain collapsed until the pressure of the blood within it is equal or superior to the pressure in the air surrounding it. It is thus possible to vary at will the resistance to the outflow of the blood from the arterial side. From the peripheral end of *R* the blood passes at a low pressure through a spiral warmer into a glass reservoir. From the reservoir a wide india-rubber tube leads to the cannula in the superior vena cava *SVC*. This cannula is provided with a thermometer. A tube placed in the inferior vena cava and connected with a water manometer M_3 shows the pressure in the right auricle. On the recording surface we thus have a record of the arterial pressure, and of the pressure within the right auricle The rate of inflow from the reservoir into the right auricle is controlled by an adjustable screw-clamp. The output through the aorta can be measured at any time by opening the tube *X*, clamping *Y*, and allowing the blood to flow for a given number of seconds. into a measure. The course of the circulation is shown in Fig. 373.

This method, although of considerable value in giving information as to the conditions which determine the output of the left ventricle and the maximum capacity of the heart as a pump, tells us nothing as to the output in the intact animal. The output as measured represents the ventricular output minus the blood flow through the coronary arteries. It is possible, however, to insert a cannula into the coronary sinus, and so to measure the blood flow through the heart muscle.* The coronary circulation must be added to the flow through the arterial resistance in order to arrive at the correct total output of the left ventricle. An essential in this preparation is that the output from the left ventricle and the pressure in the aorta must be sufficient to maintain a circulation through the coronary vessels.

Output and Inflow. It is self-evident that, provided the venous inflow remains constant, the minute output over a reasonable interval of time must also be constant, equal to the inflow, and unaffected by considerable alterations of arterial resistance or heart rate. Thus, in the heart-lung preparation with a moderate and constant venous inflow the output remains constant whether we maintain the average arterial pressure at 60 mm. Hg or at 160 mm. Hg. It is also unaffected by altering the rate of the heart from 80 beats per minute up to 160 or even 200 beats per minute. On the other hand, the output must at once be altered by changes in the venous inflow and can in fact be varied in a dog's heart weighing 50 gms. from 100 c.c. up to 3,000 c.c. per minute.

With increasing inflow of blood into the heart, the auricles and ventricles become more filled during diastole ; and during systole of the ventricles the incoming blood must accumulate in the big veins and auricles to a still greater extent. The venous pressure in the intact animal therefore rises, with increased venous inflow. In so far as venous pressure is an index of venous inflow, we may say that the output of the heart increases with the venous pressure, so long as the heart is functionally capable of dealing with the blood it receives during diastole. And since the minute output in the isolated heart is practically independent of the frequency of the beat, it follows that with a constant venous inflow the venous pressure tends to fall, and diastolic volume to diminish, as the heart rate increases. The optimum venous pressure is that which fills the ventricle during its diastole to the maximum extent from which it is able to empty. As the rate of the heart increases, the inflow of blood can also be increased without causing over-distension of the ventricles. An increase of heart rate must therefore be, in the intact animal, an important factor in enabling this organ to deal with the maximum amount of blood. Although increase of rate cannot alter the minute output with constant venous inflow, it does increase the *maximum* inflow which the heart is able to expel. At the same time, the lowering of pressure in the great veins aids the venous inflow into the chest.

As might be expected, the alterations in the vigour of the circulation depend in the first instance on the venous flow. The greater the volume of the blood that is brought up to the heart by the accessory factors of the circulation, the greater will be the output of this organ. The changes in rate and force of the heart which accompany its increased activity and increased output, *e.g.* during exercise, represent merely the response by which this organ is able to deal in the most advantageous manner with the increased inflow.

The cardiac output, like the oxygen consumption of the individual, is definitely increased after ingestion of food (Fig. 374). It is also greater in man by 10 or 12 per cent. when lying supine, than when sitting, and becomes still smaller when standing erect.*

---

* EVANS and STARLING. *J. Physiol.*, 1913, **46**, 413.

The Work of the Heart. The energy of the ventricular contraction is expended in two ways: first, in forcing a certain amount of blood into the already distended aorta against the resistance presented by the arterial blood pressure; and secondly, in imparting a certain velocity to the mass of blood so thrown out. Thus the energy of the muscular contraction is converted partly into potential energy in the form of increased distension of the arterial wall and partly into the kinetic energy represented by the momentum of the moving column of blood. The work done at each beat by each ventricle may be calculated from the formula:

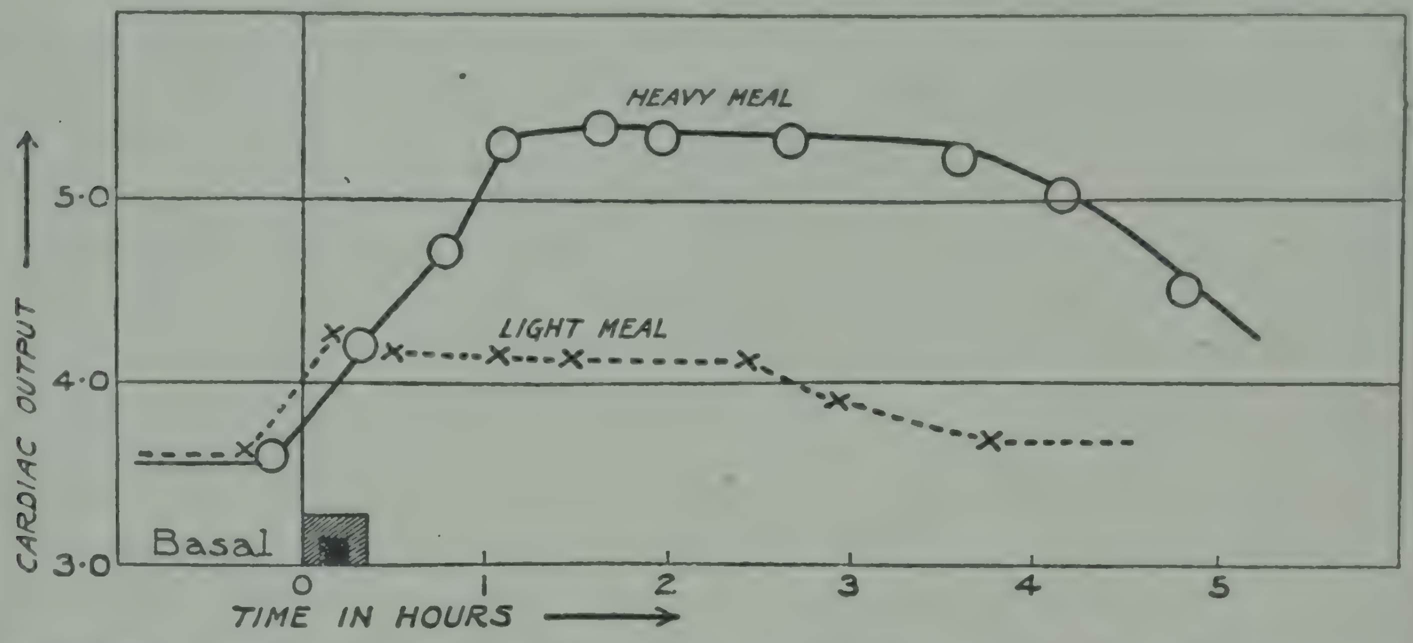


FIG. 374. Effect of food on cardiac output. (GROLLMAN.) (From MARSHALL.)

$$W = QR + \frac{wV^2}{2g}$$

where W stands for work, w for the weight, and Q for the volume of blood expelled at each contraction; R is the average arterial resistance in terms of height of a column of blood, during the outflow of blood from the heart, and V is the mean velocity at which the blood is expelled. In this equation QR is the work done in overcoming the resistance, and $\frac{wV^2}{2g}$ the energy expended in imparting velocity to the blood.

If we take 60 c.c. as the average output of each ventricle, 100 mm. Hg as the average pressure at the beginning of the aorta, and 500 mm. per second as the velocity imparted to the blood thrown into the aorta, we can roughly calculate the work done by the human left ventricle at each beat :—

$$QR = 60 \times 0{\cdot}100 \text{ m.} \times 13{\cdot}6 = 81{\cdot}6 \text{ gram-metres,}$$

or roughly 80 gram-metres. On the other hand, the expression

$$\frac{wW^2}{2g} = \frac{60 \times (0{\cdot}5)^2}{2 \times 9{\cdot}8} = 0{\cdot}7 \text{ gram-metres.}$$

It is evident that this latter factor is negligible, and that with small outputs we may regard the work of the heart as sufficiently expressed by the output multiplied by the average arterial blood pressure. Taking the average pressure in the pulmonary artery at 20 mm. Hg, the work of the right ventricle at each beat would amount to about 16 gram-metres, a total for the two ventricles of about 100 gram-metres per beat, which is equivalent to about 10,000 kilogram-metres in twenty-four hours for a man at rest.

During muscular work this figure would be largely increased. Not only does QR become much larger, but the velocity factor is no longer negligible, since the work done in imparting velocity to the blood increases as the cube of the output per minute. If we take, as an example, a maximum effort on the part of an athlete, we may assume an output per beat of 180 c.c. and a pulse rate of 180 per minute (an output per minute of 32·4 litres) and an average arterial pressure of 120 mm. Hg.

---

* DONAL, GAMBLE and SHAW. *Amer. J. Physiol.*, 1934, **109**, 666; MCMICHAEL. *Q. J. Exp. Physiol.*, 1937, **27**, 55.

Then

$$QR = 180 \times 0{\cdot}120 \times 13{\cdot}6 = 294 \text{ gram-metres.}$$

To determine the velocity of output, we assume that 180 c.c. of blood are thrown out into the aorta during $\frac{1}{8}$ second. This gives a velocity of 2·3 metres per second, assuming a cross section of 625 mm.^{2} at the root of the aorta. Therefore

$$\frac{wV^2}{2g} = \frac{180 \times (2{\cdot}3)^2}{2 \times 9{\cdot}8} = 49 \text{ gram-metres.}$$

The total work of both sides of the heart will be :

$$\underset{\text{Left side.}}{294 + 49} + \underset{\text{Right side.}}{59 + 49} = 451 \text{ gram-metres per beat, or 81 kilogram-metres per minute.}$$

It will be noted that now the velocity-factor accounts for nearly one-quarter of the total work.

Since the period of systolic output occupies not more than $\frac{3}{8}$ of the cardiac cycle, the mean velocity during expulsion must be $2\frac{2}{3}$ times greater than the mean aortic velocity : the following general formula $W = \frac{7QR}{6} + \frac{w(VC)^2}{gE^2}$, (where V is the mean aortic velocity, C the duration of the cardiac cycle and E the duration of period of expulsion) gives the work of both ventricles.* The expression QR is also only approximately correct. Supposing the pressure in the aorta at the beginning of systole is 50 mm. Hg, and at the end of systole 150 mm., the work could not be deduced accurately from the average pressure, but would need a simple application of the integral calculus for its determination. This has been made by Katz,† who finds that the expression employed above deviates from the real value by a variable amount, and tends to underestimate the work done, especially that due to the kinetic factor.

It is important to remember that the strain or tension, which is thrown on the cardiac fibres and which resists their contraction, will be determined not only by the blood pressure which has to be overcome, but also by the size of the ventricle cavities. If we regard the ventricle as a sphere, it is evident that, in order to produce a pressure p per unit area, in its interior, the total pressure exerted by the walls must be $4\pi r^2 p$, *i.e.* it must be proportional to the surface area of the inside of the ventricle or to the squareof its radius. Moreover, with increasing distension, the wall of the heart becomes thinner and the number of muscle fibres in a given area fewer, so that the larger the heart the more strongly will each fibre have to contract in order to produce a given tension in the contained fluid. At the beginning of systole the distended heart must therefore contract more strongly than at the end of the systole, if it is to develop an endocardiac pressure sufficient to overcome that in the aorta.

Adjustment of Output to Inflow. The heart in the heart-lung preparation has a marvellous power of regulating its activity according to the mechanical demands which are made upon it. Thus, if we keep the venous inflow constant, it makes no difference to the output of the heart whether the average arterial pressure be maintained at 80 or 160 mm. Hg, although in the latter case the heart must do twice as much work in order to keep the outflow at the same level. We could compare this behaviour with that of an automobile with an automatic throttle designed to give a constant velocity, whatever the load or the gradient. Again, if we maintain the arterial pressure constant and alter the venous inflow, we find that within very wide limits the heart is able to expel against the arterial resistance the whole of the blood which flows into it from the veins. In this way we can alter the output of a small heart of 50 grammes from 300 to 3,000 c.c. per minute. To use our automobile analogy again, the throttle is now designed to enable it to perform any reasonable journey, long or short, in the same time.

---

* Lovatt Evans. *J. Physiol.*, 1918, **52**, 6.

† Katz. *Amer. J. Physiol.*, 1932, **99**, 579.

Oxygen Usage and Efficiency. As we should expect, this variation in the work done by the heart is associated with corresponding variations in the chemical changes which occur at each heart beat. It is easy with the heart-lung preparation to determine the actual work done and to vary this at will. Then, knowing that each cubic centimetre of oxygen used represents the liberation of about 2 kg.-m. of energy, it is easy to calculate at what mechanical efficiency the heart is working over any given period of observation. Thus, if a heart uses 3,000 c.c. of oxygen in a given period it has liberated as heat and work about 6,000 kg.-m. of energy; if in that time it performed 600 kg.-m. of work the efficiency was 0·1 (or 10 per cent.). In this way it was found that the efficiency of the heart is probably never greater than 30 per cent.—more usually it is from 5 to 10 per cent. In this respect the heart shows much the same mechanical performance as the skeletal musculature. The respiratory exchange of the heart increases *pari passu* with the work it has to do.* Thus in an isolated dog's heart, weighing 70 grammes, with a constant output of 35 litres per hour, raising the arterial pressure from 80 mm. Hg to 140 mm. Hg increased the oxygen consumption from 228 to 404 c.c. per hour. In another experiment with a heart of 59 grammes, in which the arterial pressure was maintained constant at 80 mm. Hg, increasing the output from 9·3 to 92 litres per hour raised the oxygen consumption from 155 to 649 c.c. per hour. An increase of work effected by a rise of arterial resistance is attended by a greater increase of oxygen usage than is the same increase of work effected by a rise of output. When the efficiency of the heart at various levels of work is plotted against the mechanical work done, it is seen to reach a maximum with increasing work, and then to decline (Fig. 375). Increase of the work beyond this point soon leads to insufficiency: over a certain further range more work can be accomplished but at rapidly increasing cost, and with still further increase of resistance or inflow the mechanical work done also declines and failure is then imminent.

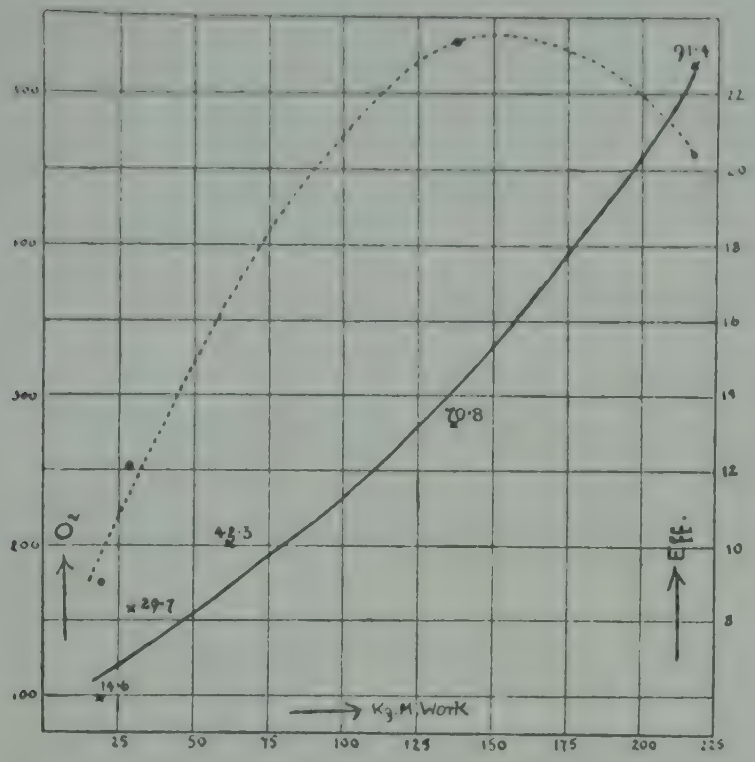


FIG. 375. Oxygen Usage and Efficiency of Dog's Heart (H–L–P), showing their Relation to Work done. Arterial pressure, 80 mm. Hg. throughout. Continuous line, oxygen usage, c.c. p. hr: the figures on the oxygen curve show the output in l. p. hr. Dotted line, efficiency. (LOVATT EVANS, 1918.)

The Law of the Heart. Investigation of the volume and pressure changes of the heart under varying conditions of arterial resistance and venous filling throws light on the mechanism of the adaptation.† Let us take first the changes in volume as recorded by the cardiometer, and suppose a heart is contracting 100 times per minute, forcing out at each beat 10 c.c. of blood into the aorta against an average pressure of 100 mm. Hg, with systolic and diastolic pressures respectively of 110 and 90 mm. Hg (Fig. 376). The aortic valves will open as soon as the pressure rises above 90 mm. Hg. In order that the left ventricle may force out 10 c.c. of blood the pressure in its interior must rise at each heart beat above the maximum systolic pressure in the aorta, *e.g.* to 120 mm. Hg. Suppose the arterial resistance is now increased so as to bring the average pressure up to 130 mm. Hg. The heart now may raise the pressure in its interior to 130 mm. Hg. This will be higher than the diastolic pressure in the aorta and a certain amount of blood will escape, but the outflow of blood will cease as soon as the pressure in the aorta is equal to that in the ventricle. Diastole will then occur, the ventricle will relax before it has emptied out 10 c.c. of blood. Let us assume it has forced out 3 c.c. of blood, so that the excess of 7 c.c. remains at the end of systole. Meanwhile, the venous inflow is proceeding at the same rate as before, so that at the end of diastole its diastolic volume will be increased by 7 c.c. At the next beat we find that the contraction of the ventricle is much more forcible. The maximum pressure now rises to 140 mm. Hg and 8 c.c. of blood are sent out into the aorta. At the end of this beat the heart will be still fuller than before, containing an excess of 9 c.c. of blood. The third beat is still more forcible, the intraventricular pressure rising to a maximum of 145 mm. Hg, and now 10 c.c. of blood are expelled. After this, the heart goes on beating regularly, expelling 10 c.c. of blood at each beat, *i.e.* the same amount as it receives from the veins. But the heart remains more dilated than it was previously (Fig. 376), since it contains an excess of 9 c.c. of blood. If now the arterial resistance be suddenly reduced to its previous amount, the first beat after the change may send out 17 c.c. of blood, the second beat 12 c.c. of blood and the third beat 10 c.c. as before. It is similar with alterations of the rate of inflow, with constant arterial pressure, as shown in Fig. 377. We see, therefore, that an increased liberation of energy results from the contraction of the heart muscle when there is an increased diastolic volume of the heart and *vice versâ* ; but increased diastolic volume means greater length of the muscular fibres composing its wall, so we arrive at a statement that, as in voluntary muscles, the *energy of contraction is a function of the length of the muscle fibres.* This important generalisation is commonly called Starling's "Law of the Heart." †

---

* EVANS and MATSUOKA. *J. Physiol.*, 1914, **49,** 378.

† PATTERSON and STARLING. *J. Physiol.*, 1914, **48,** 357; PATTERSON, PIPER and STARLING, *ibid.*, p. 465.

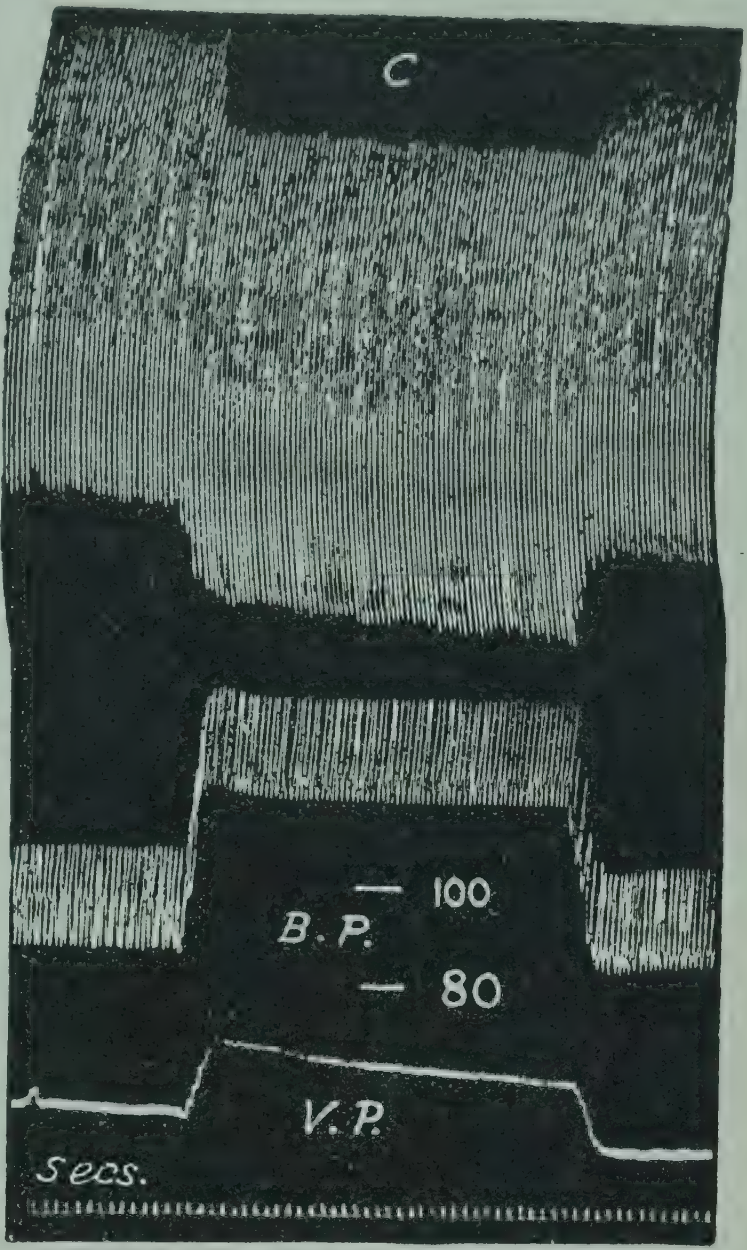


Fig. 376. Effect of Increased Arterial Pressure on the Volume Changes of the Heart, with a steady inflow of 154 c.c. Blood per 10 seconds.

C. = cardiometer curve; downstroke is diastole. BP = arterial blood pressure. V.P. = pressure in the inferior vena cava. The lines 100 and 80 are calibrations for the blood pressure in mm. Hg. (Patterson, Piper and Starling.)*

In a resting heart increasing distension of its cavities will tend to stretch its muscle fibres and therefore to exert a tension on them. By accurately recording the pressure changes within the contracting ventricle under varying conditions, it has been found that the tension on the fibres is not the determining factor. In a heart beating regularly, the inflow of blood is proceeding during the relaxation of the ventricles, *i.e.* the muscles are giving before the inflowing blood. The latter is therefore able to distend the heart without exercising more than a minimum pressure on its walls. With a small inflow it is found that the pressure in the ventricles may be approximately zero at the end of diastole, whether the heart is contracting against a resistance of 80 mm. Hg or a resistance of 120 mm. Hg, or, with a moderate arterial resistance, whether it is receiving 5 c.c. or 0 c.c. during the period of diastole. With a still larger outflow or a bigger resistance there is an appreciable rise of tension within the ventricles at the end of diastole. But that the change in *distension*, *i.e.* in the diastolic length of the muscle fibres, is the chief factor in determining the energy of the subsequent contraction is evident on reference to Fig. 378 which shows the reaction to increasing filling of the tortoise ventricle contracting isometrically. It will be seen that as the diastolic filling increased from 0 to 1·5 c.c., the intraventricular pressure developed during systole increased from 0 to 45 mm. Hg. The maximum systolic pressure (48 mm. Hg) was developed with a filling of 3 c.c. and an initial tension within the heart of 3 mm. Hg. Further increasing the filling to 4 c.c. caused for the first time a considerable rise of initial tension to 13 mm. Hg, but the result was to *diminish* the tension developed during systole. This means that, up to a certain optimum length, the energy of contraction increases with every increase in initial length of the fibres. Beyond this point it requires considerable tension to cause a further increase in length, but this is associated with a diminution in the effectiveness of the contractions.

---

* Patterson, Piper and Starling. *J. Physiol.*, 1914, **48**, 465.

† Starling. Linacre Lecture, "The Law of the Heart." (London, 1918.)

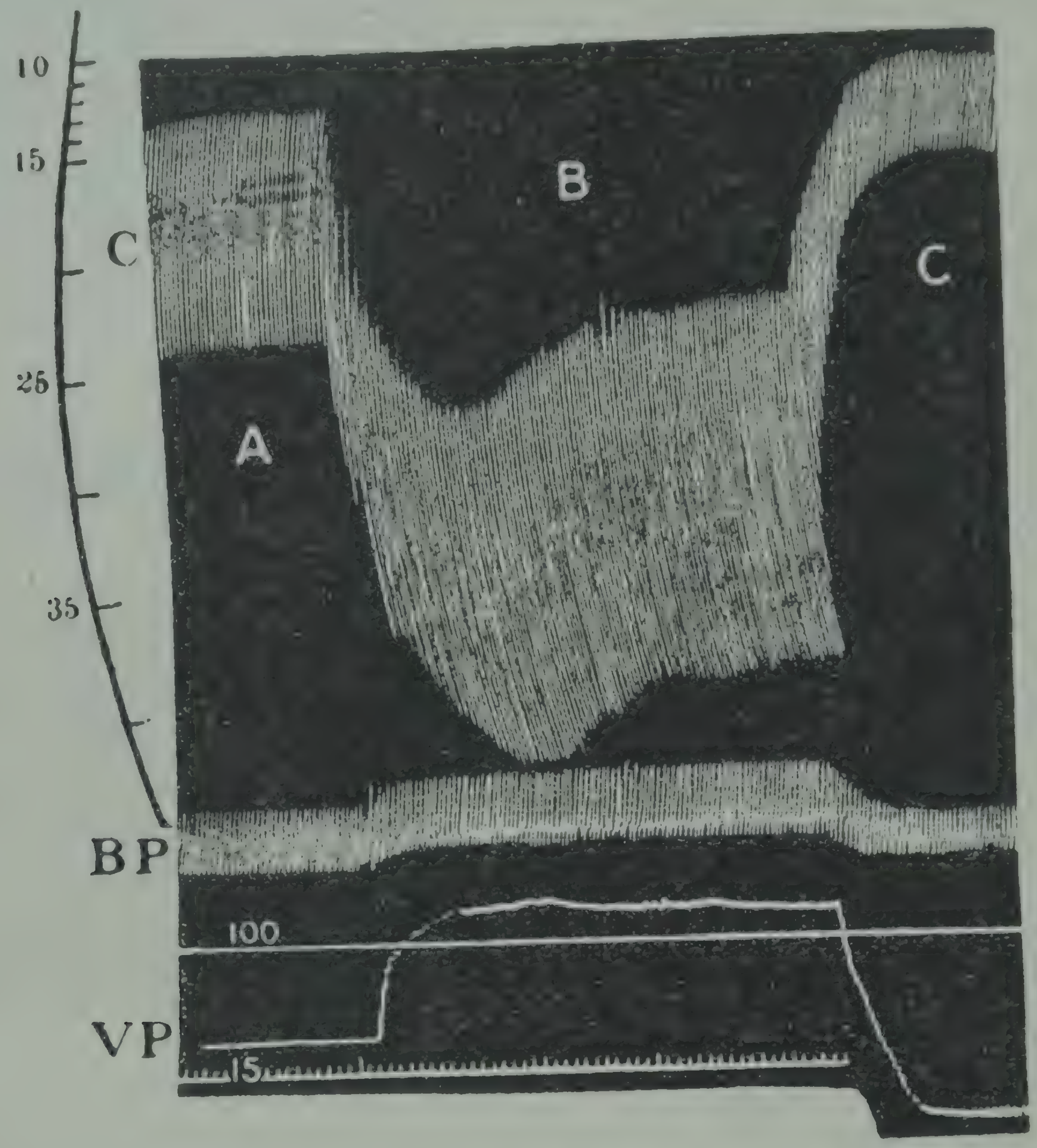


Fig. 377. Effect of Alterations in Venous Supply on Volume of Heart. Heart 67 grms.

	Arterial pressure. mm.Hg.	Venous pressure. mm. H_2O.	Output of heart in 10 secs. cc.
A	124	95	86
B	130	145	140
C	124	55	33

The curved line at the side represents the value of the cardiometer excursions, *i.e.* of alterations of ventricular volumes, in cubic centimetres. (Patterson, Piper and Starling, 1914.)

The capacity of the heart for adapting itself to changes in mechanical demands made upon it will be limited by its inability to dilate further; in the intact animal its dilatation is limited by the pericardium, and by the mechanical disadvantage at which the further dilated heart acts. The greater its diastolic volume, and the more nearly globular its form, the greater the mechanical disadvantage of the muscle fibres in raising the pressure in the interior of the ventricles, so that by continually increasing the demands on the heart, we shall finally arrive at a stage at which it is unable to deal with the blood supplied to it.

Evidence of a relation between heart volume and total energy consumption was furnished by the demonstration, by Evans and Matsuoka, that there is a parallel between heart volume and oxygen usage. Determinations of the oxygen usage of the heart * also show that, within limits, when the diastolic volume is the same, the oxygen usage is the same, whatever may be the amount of work done (Fig. 379): in other words, the oxygen utilisation and the diastolic volume are closely correlated over rather wide ranges.‡ Beyond the point of maximal performance diastolic volume and oxygen utilisation show a further increase, but the work accomplished diminishes; this holds good in general whenever, with an increased volume in diastole, there is a reduced efficiency, as is the case when the heart fails. This was confirmed by Stella,§ who used the tortoise heart and found that the relationship between diastolic volume and oxygen usage depends upon the conditions under which the heart works.

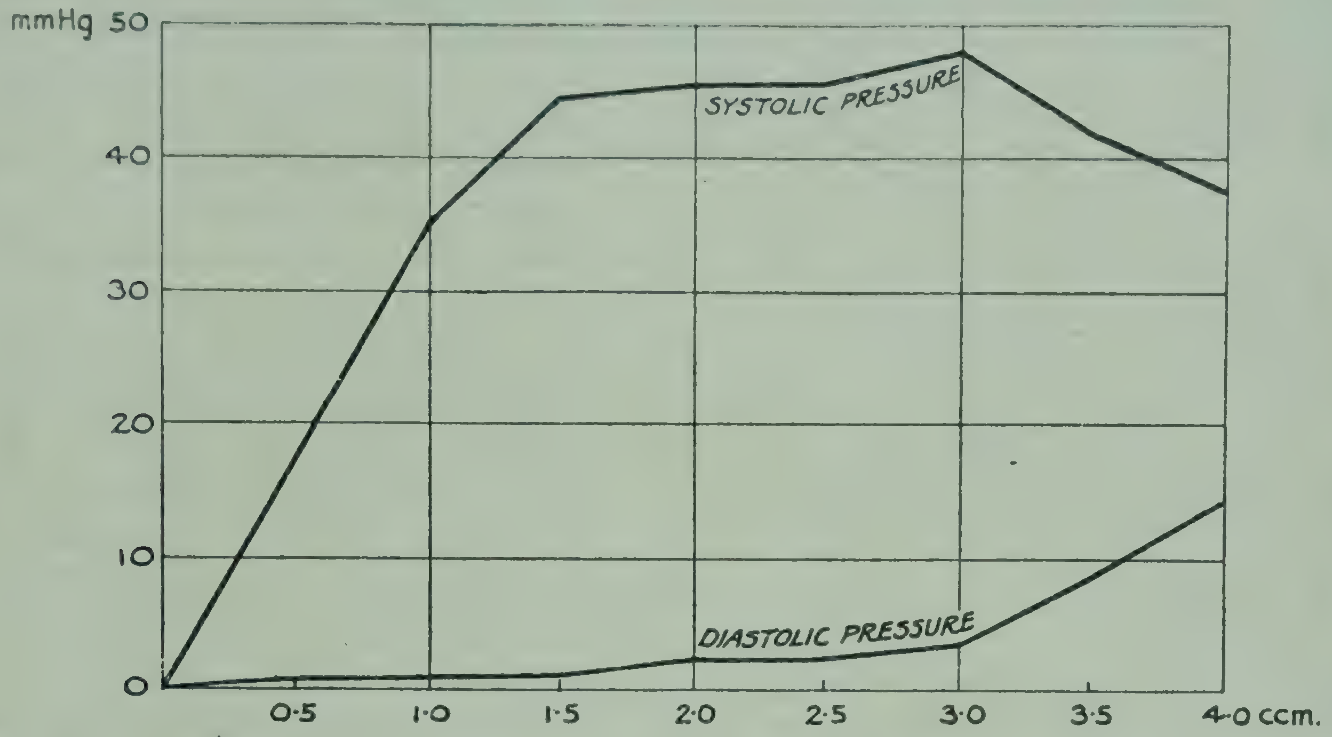


FIG. 378. Relation of the Tension developed during Isometric Contraction of the Tortoise Ventricle to the Filling (*i.e.* initial length of fibres) of the Ventricle. (KOZAWA.)†

The physiological condition of the heart is measured by the maximum pressure which it is able to produce in its cavities when it contracts, starting from a certain initial size or length of fibres. As the heart becomes fatigued this pressure falls, so that the heart must dilate in order that each contraction shall produce the same maximum pressure as before. Dilatation is therefore a measure of fatigue.

THE HYPODYNAMIC OR INSUFFICIENT HEART is unable properly to cope with the work which it is set to do. Sooner or later all isolated hearts become insufficient, a condition which is revealed by a rising intra-auricular pressure. As the heart approaches insufficiency it tends to do less work, and to dilate. The dilatation is always accompanied by an increased oxygen usage. If now the work conditions are lightened the dilatation and oxygen consumption may be reduced, and the heart may be competent to carry on at the easier level. If, on the contrary, still further work is offered to an already insufficient heart the mechanical work done *may* be restored to its former level, but only at the cost of a still further dilatation, rise of intra-auricular pressure, and greatly augmented oxygen usage. Like all pathological phenomena, these effects are nothing more than exaggerations of normal properties, and insufficiency is mainly a reduced mechanical efficiency; of the expended chemical energy (as judged by oxygen usage) at a given diastolic volume, a smaller fraction is converted into mechanical work by the failing than by the normal heart.

Addition of digitalis or strophanthin to the blood leads to a real improvement of the insufficient heart, *i.e.* at a given task the diastolic volume and oxygen consumption of the spontaneously insufficient heart are reduced and hence the efficiency is increased.

TONE. Many authors speak of a "tone" of the heart muscle, generally meaning thereby a tendency to resist distension during diastole. This term is ambiguous, and has been used in so many meanings that it would be better to avoid the word altogether. The "tone of the heart" is really synonymous with physiological condition. A heart in good condition has a "high tone." It empties itself almost completely at each beat, even when receiving a considerable quantity of blood during diastole. A heart with a "low tone" is in the condition of a fatigued heart. It is greatly dilated and at the end of systole still contains a large amount of residual blood.

---

* VISSCHER and STARLING. *J. Physiol.*, 1927, **62**, 243.

† KOZAWA. *J. Physiol.*, 1915, **49**, 233.

‡ HEMINGWAY and FEE. *J. Physiol.*, 1927, **63**, 299.

§ STELLA. *J. Physiol.*, 1931, **72**, 247.

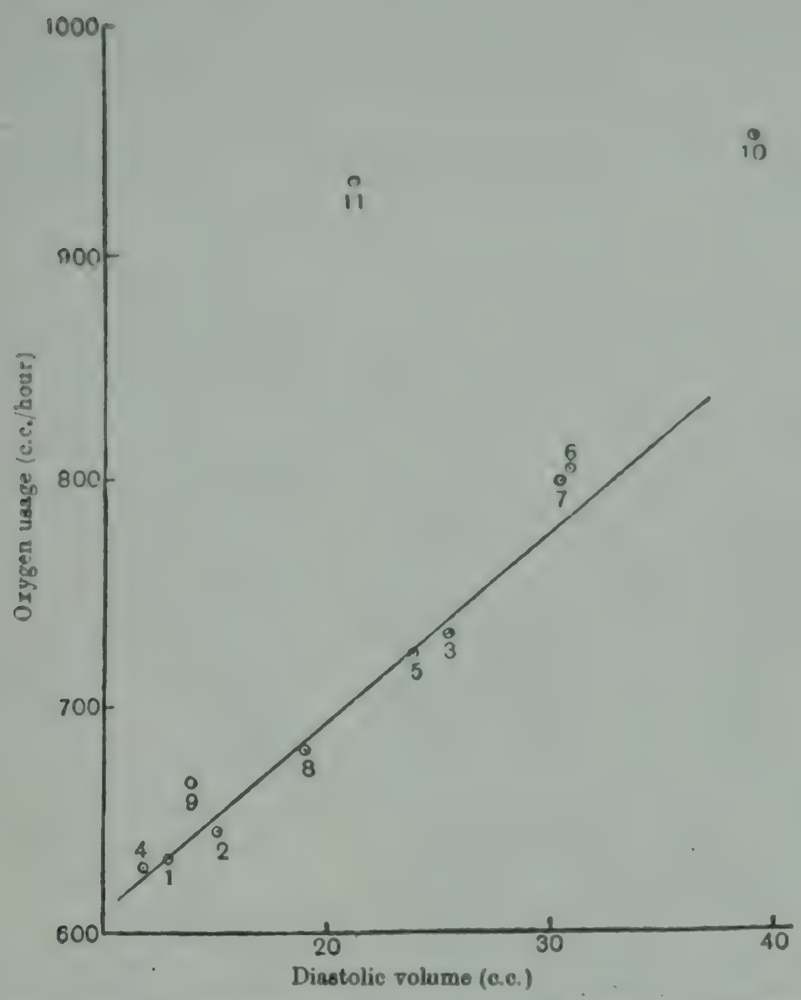


FIG. 379. Showing Relation between Heart Diastolic Volume and Oxygen Usage of whole heart-lung preparation. (HEMINGWAY and FEE. *J. Physiol.*, 1927, **63**, 299.)

COMPENSATION. This property of the cardiac muscle is responsible for the power of "compensation" possessed by a diseased heart. We may take as an example the destruction of one aortic valve, a lesion which can be produced experimentally in a dog. Immediately after the lesion is established, no additional resistance is offered to the expulsion of the blood, and the ventricle will send the normal amount into the aorta. During the succeeding diastole the blood at a high pressure in the aorta will leak back into the ventricle through the damaged valve. The arterial pressure therefore falls rapidly, and the ventricle receives blood from two sides, *i.e.* by regurgitation through the aortic valves, and in the normal way from the auricles and veins. At the end of diastole the ventricle is therefore overfilled. Increased stretching of its fibres, however, has the effect of producing an increased contraction, and the heart, at its next systole, throws out not only the normal quantity of blood, but also that which it has received back from the aorta. The arterial system thus receives at each beat the normal quantity of blood *plus* the amount which leaks back into the ventricle after each systole; so that the amount of blood remaining in the aorta and available for passage on to the capillaries is the same as in the normal animal. On this account, after a lesion of the aortic valves has been established, the mean arterial pressure remains the same as before, although the pulse pressure at each heart beat is increased in amplitude. The augmented output by the ventricles naturally involves increased work on the part of their muscular walls, which react to increased work by hypertrophy. The final effect, therefore, is a heart bigger than normal, with hypertrophied and thickened walls, but capable of maintaining an adequate circulation throughout all parts of the body; in other words, complete compensation has taken place.

For Reference

MEEK (1927). The Question of Cardiac Tonus. *Physiol. Rev.*, **7**, 259.

Influence of Temperature on the Heart. Warming the heart quickens the rhythm, and cold slows it. In the isolated heart beating in normal sequence, the shortening of the contractions of the different cavities with rise of temperature may be ascribed to the quickening of the sinus rhythm. The changes, however, *i.e.* the shortening of the contraction with rise and the lengthening with fall of temperature, are more marked than when the rhythm is changed by direct excitation of the sino-auricular node itself. The increase of rhythm with temperature is almost proportional to the temperature, and over a limited range this proportionality may be linear.* At 40° C. the isolated mammalian heart may beat four times as frequently as 25° C. Warming the heart beyond 44° C. may cause stoppage of the heart or send it into fibrillation.

---

* KNOWLTON and STARLING. *J. Physiol.*, 1912, **44**, 206.

Similar effects are produced on conduction, both in the heart muscle itself and in the special conducting tissues. Thus, cooling of the heart leads to an increase in the A.V. interval.

The contraction of cardiac muscle is affected in the same way, a rise of temperature lessening the duration both of the contraction and of the refractory period.

As in a skeletal muscle, the slow contraction of the heart at low temperatures is more efficient than the quick contraction at high temperatures, *i.e.* when the muscle contracts slowly and maintains the intraventricular tension for a longer time, less energy is required for the expulsion of a certain volume of blood than at a higher temperature.

Influence of the Composition of the Surrounding Medium on the Heart Muscle. All the cells of the body require a definite osmotic environment. This is equivalent to a 0·7 per cent. sodium chloride solution for the frog's heart, and to a 0·9 per cent. solution for the mammalian heart. As Ringer first showed, the nature of the salts employed for making up the isotonic solution is all-important to heart muscle.* Thus, a strip of muscle from the apex of the tortoise's ventricle immersed in a 0·7 per cent. solution of sodium chloride, begins to beat rhythmically. The contractions soon reach a maximum and then die away in a condition of relaxation. Sodium chloride therefore acts as a stimulus to contraction, but is unable to maintain the beats for any considerable length of time. On now adding to the solution a trace of calcium chloride the contractions begin again (Fig. 380). The relaxations after each contraction then become more and more incomplete, until finally the heart

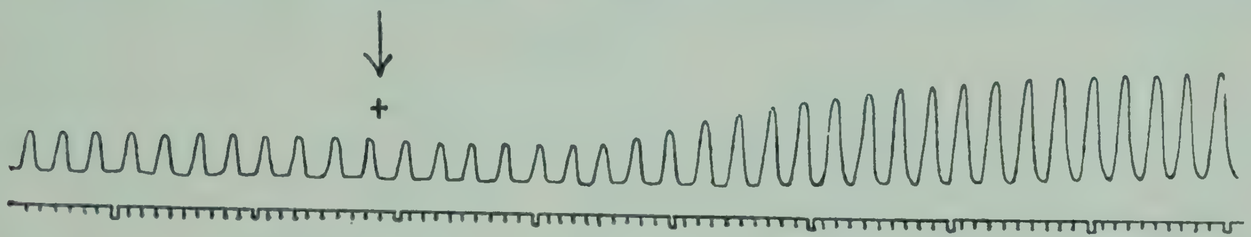


FIG. 380. Tracing of Contractions of a Frog's Heart, showing effect of adding a trace of $CaCl_2$ to the NaCl Solution used previously for Perfusion. (RINGER.)

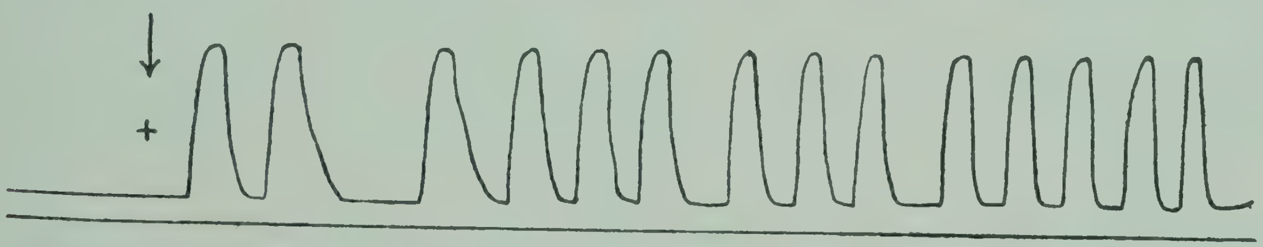


FIG. 381. A Frog's Heart, Poisoned by Excess of Calcium Salts, Recovers its Spontaneous Rhythm on adding a trace of KCl to the Perfusion Fluid. (RINGER.)

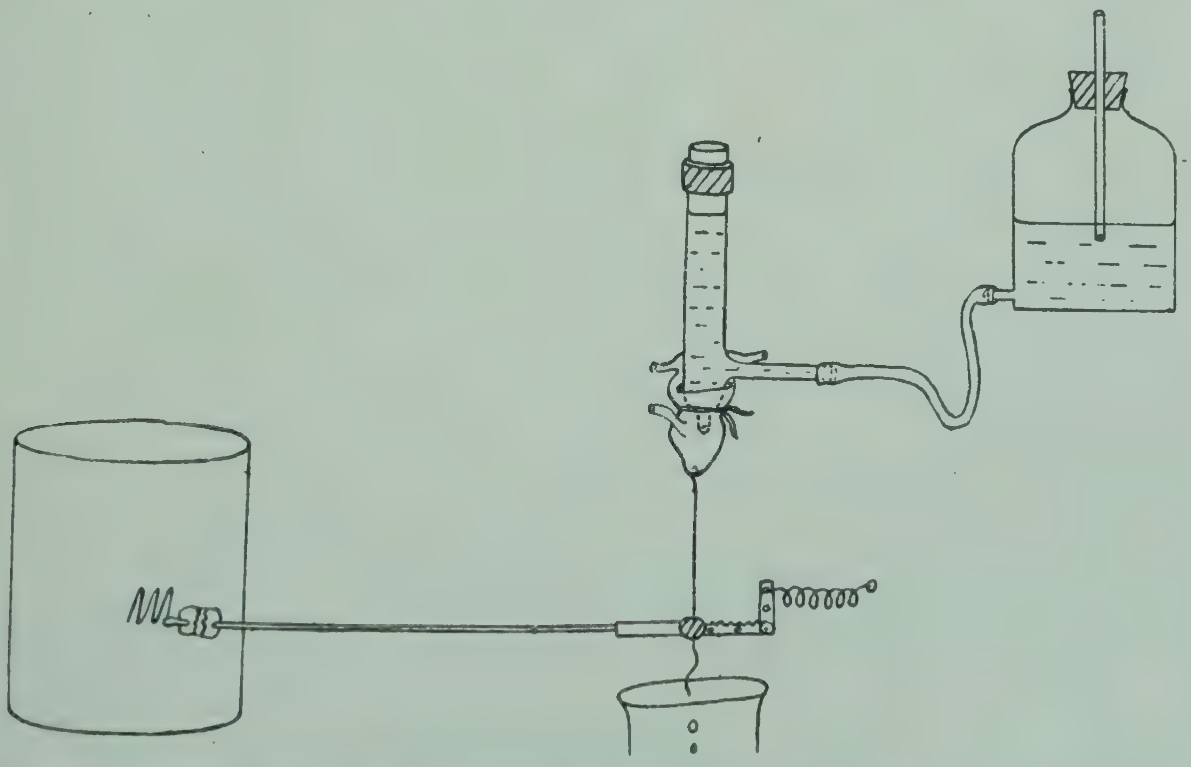


FIG. 382. Symes' Apparatus for Perfusion of Frog's Heart. (HARRIS.)

[*] RINGER. *J. Physiol.*, 1883, **4**, 29.

stops in a tonically contracted condition. If now a trace of potassium chloride be added, the contractions recommence and may last for many hours, although the solution contains nothing which can furnish energy to the contracting muscle (Fig. 381).

The exact significance of the different ions for the functions of cardiac and other forms of muscular tissue must still be regarded as an open question.

Fluids containing the three ions mentioned above in slightly varying proportions, are commonly used to maintain the beat in an excised heart either of a cold- or of a warm-blooded animal (Ringer's fluid, etc., *v.* p. 62). In the case of the latter, it is necessary to keep the fluid warm, and saturated with oxygen. Locke* showed that the addition of glucose to the solutions enables the beats to go on for a longer period of time, and will in fact renew the rhythm of a heart which has ceased beating while being fed with pure saline solution. Addition of sodium lactate is even better.

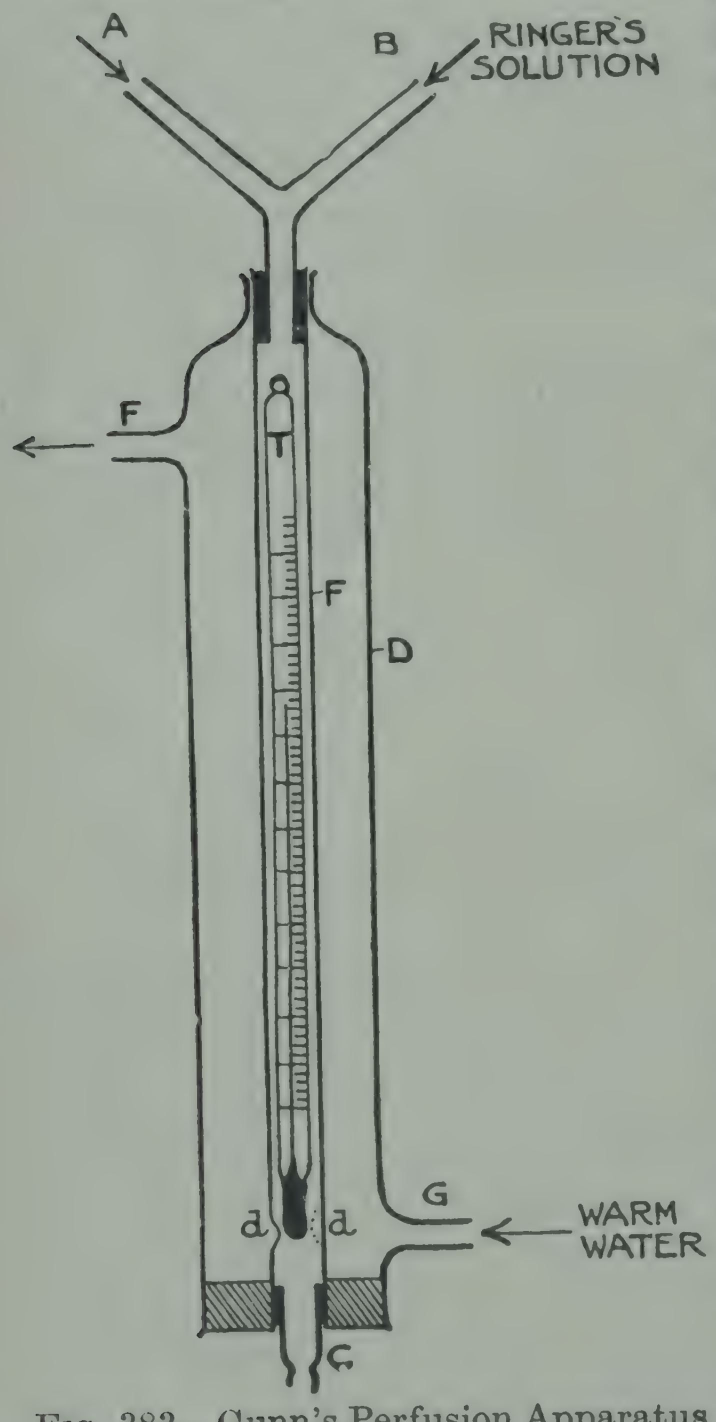


FIG. 383. Gunn's Perfusion Apparatus for the Mammalian Heart.

The influence of the chemical composition of the medium on the contraction of the heart may be investigated in the following ways:

One of the simplest methods of investigating the heart of cold-blooded animals is by perfusing the heart cavities with the fluid under investigation.

In one method, the fluid is allowed to flow through the whole heart, passing in by the sinus and out by the aortæ, or, as in Symes' method,† in by a cannula tied in the auricle, and out by the aortæ (Fig. 382). It is usual to start with Ringer's fluid, and then to omit or increase the constituents one by one, in order to demonstrate their influence.

The heart of warm-blooded animals can also be investigated by a somewhat similar method. It was shown by Porter ‡ that the mammalian heart could be kept alive by perfusing oxygenated blood serum through the coronary vessels, and Locke § found that the same results could be obtained by using warm oxygenated Ringer's solution of the same tonicity as mammalian blood. A simple apparatus for this purpose is shown in Fig 382.‖ It consists of an outer jacket, D, through which a constant stream of water at a little above body temperature is passed. The inner tube, F, is made of thin glass, and at its lower end three small indentations, d, d, serve to support a thermometer, T, of such a thickness that only a small space is left between it and the inner tube. Into the lower end of the inner tube the cannula, C, is fixed by a rubber sleeve, while at the upper end a Y-shaped tube delivers either oxygenated Ringer's solution by B, or any other solution which is to be perfused by A. The proper perfusion pressure is provided by placing the reservoir of Ringer's, or other, solution at the desired height, and these reservoirs are immersed in a bath kept at body temperature. The heart having been excised and washed well in cold saline, the cannula is tied into the stump of the aorta, and the perfusion started. The pressure of the fluid closes the aortic valves, so that the solution only passes into the coronary vessels. After passing through the heart, it reaches the right side, and flows away. A thread is attached to the apex of the heart, and the beats are recorded by means of a lever attached to the thread.

HYDROGEN ION CONCENTRATION. The reaction of the perfusing fluid has considerable influence on the frog's heart. When weak acids are perfused through this heart, it dilates, the beats become smaller, slower, and finally at about pH 2 or 3, disappear. A similar relaxation may be obtained as the result of the action of carbon dioxide. On perfusing with weak alkalies on the other hand, the relaxation during diastole becomes more and more incomplete, so that the heart is finally arrested at pH 11 or 12 * in a contracted condition. There will thus be some reaction, actually at pH 8·3,† which will represent the optimum reaction for the beat of the frog's heart. This optimum reaction is not the same for the different cavities of the heart, nor for the hearts from various animals.

---

* LOCKE. *J. Physiol.*, 1895, **18**, 333.
† SYMES. *J. Physiol.*, 1918, **62**, Proc. xlvi.
‡ PORTER. *Amer. J. Physiol.*, 1898, **1**, 141.
§ LOCKE. *J. Physiol.*, 1904, **31**, Proc. xiii.
‖ GUNN. *J Physiol.*, 1913, **46**, 506.

In the mammal under normal conditions, the chief variable factor affecting the reaction of the blood is the tension of carbon dioxide in this fluid; and any change in the carbon dioxide of the blood produces corresponding alterations in the heart's action. The effective factor is probably always the hydrogen ion concentration, so that a deficiency of CO_2 may be partially compensated for by reduction of the $NaHCO_3$, *e.g.* by an increased amount of lactic acid in the blood. The optimum pH is 7·5 to 8·0. Decreased pH (reduced alkalinity) causes (1) Slowing of the cardiac rhythm, from its action on the sino-auricular node. (2) Slower propagation and deficient spread of the excitatory wave through the auricular muscle and the A.V. bundle.‡ (3) Dilatation of the ventricles, which is greater than can be accounted for by the slowing of the heart and the greater filling thereby induced (cp. Fig. 384). (4) Dilatation of coronary vessels.

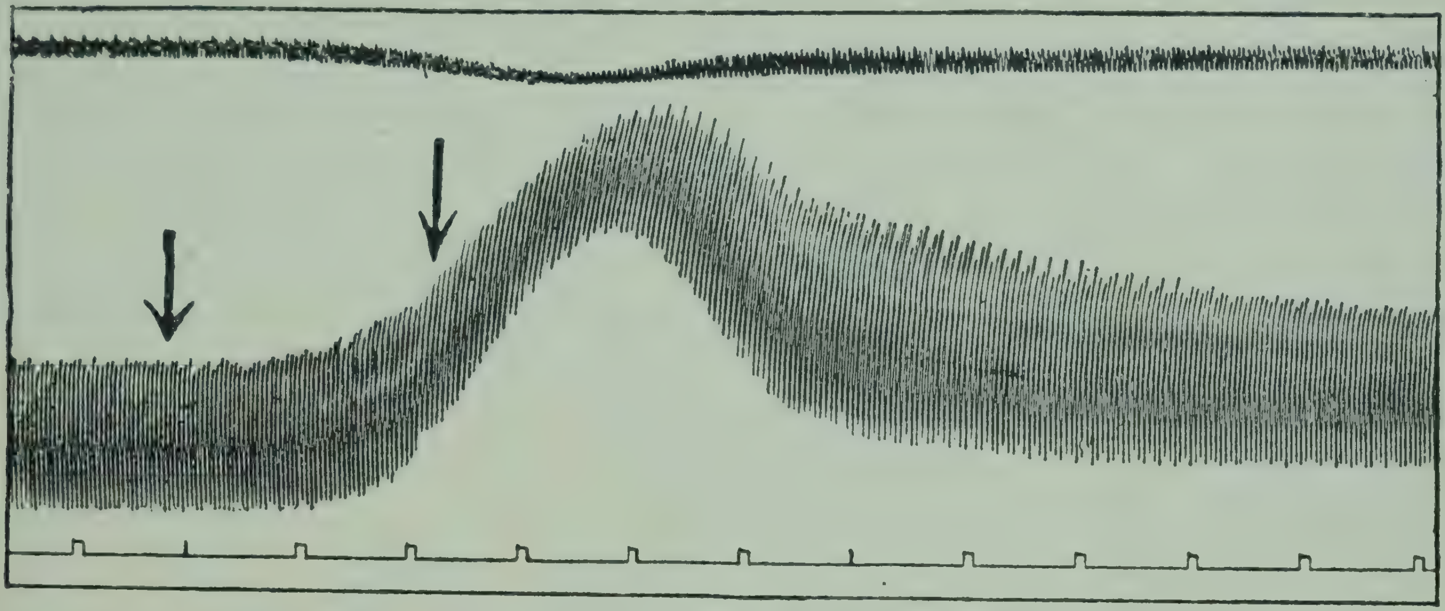


Fig. 384. Volume Curve of Ventricles (Cat) (lower curve).

The upper curve is the arterial pressure, maintained by an adjustable resistance at 130 mm. Hg. Between the arrows the air used for artificial respiration was replaced by a mixture containing 20 per cent. CO_2 and 25 per cent. oxygen. Note the dilatation with impaired contraction, followed by increased amplitude of contraction as the heart recovers. (Jerusalem and Starling, *J. Physiol.*, 1910, **40**, 279.)

Yandell Henderson § found that vigorous artificial ventilation of the lungs brought about a condition in which the heart's contraction was very forcible but its cavities almost empty. He ascribed this condition to the hypertonicity of the heart muscle produced by washing the carbon dioxide out of the blood. The condition of low CO_2 content of the blood (and consequent raised pH) is known as "*acapnia.*" Dale and Evans ‖ showed that, in the condition of acapnia there is paralysis of the spinal vasomotor centres and a large fall of arterial pressure. The inflow of blood to the heart is therefore decreased, and the heart, having only a small amount of blood to contract on, looks "hypertonic," *i.e.* reduced in size. The conditions may be at once removed by adding 5 per cent. CO_2 to the air used for the artificial ventilation. When the heart is perfused with a saline fluid of very alkaline reaction (pH 8·4) the coronary flow is so reduced, by constriction of the vessels, that the heart fails owing to oxygen lack. Up to pH 8·0 there is much less effect.

THE NERVOUS REGULATION OF THE HEART

So far we have considered the action of the heart as a separate organ, but, in order that it may be adapted to the needs of the body as a whole, its automatic mechanism must be subject to the central nervous system, which must be able, in the general interest, to regulate the heart by increasing or diminishing its activity.

---

* Mines. *J. Physiol.*, 1913, **46**, 216.

† Clark. *J. Physiol.*, 1913, **47**, 70.

‡ Drury and Andrus. *J. Physiol.*, 1924, **59**, Proc. xli.

§ Yandell Henderson. *Amer. J. Physiol.*, 1908–1910 (several papers).

‖ Dale and Evans. *J. Physiol.*, 1922, **56**, 125.

The regulation is effected through the intermediation of various afferent nerves, and of efferent nerve fibres to the heart. The afferent supply from the heart itself is in afferent fibres in the right vagus and in the aortic nerve; the efferent, the vagus and sympathetic. They are not essential to life. Thus a dog, in which all the nerves of the heart had been divided, survived the operation for eight months, the pulse rate during the time not having appreciably altered and the animal being in a fair condition of health, though less capable of carrying out work, such as running, since the mechanism for adjusting the action of the heart in response to the needs of the muscles had been lost.

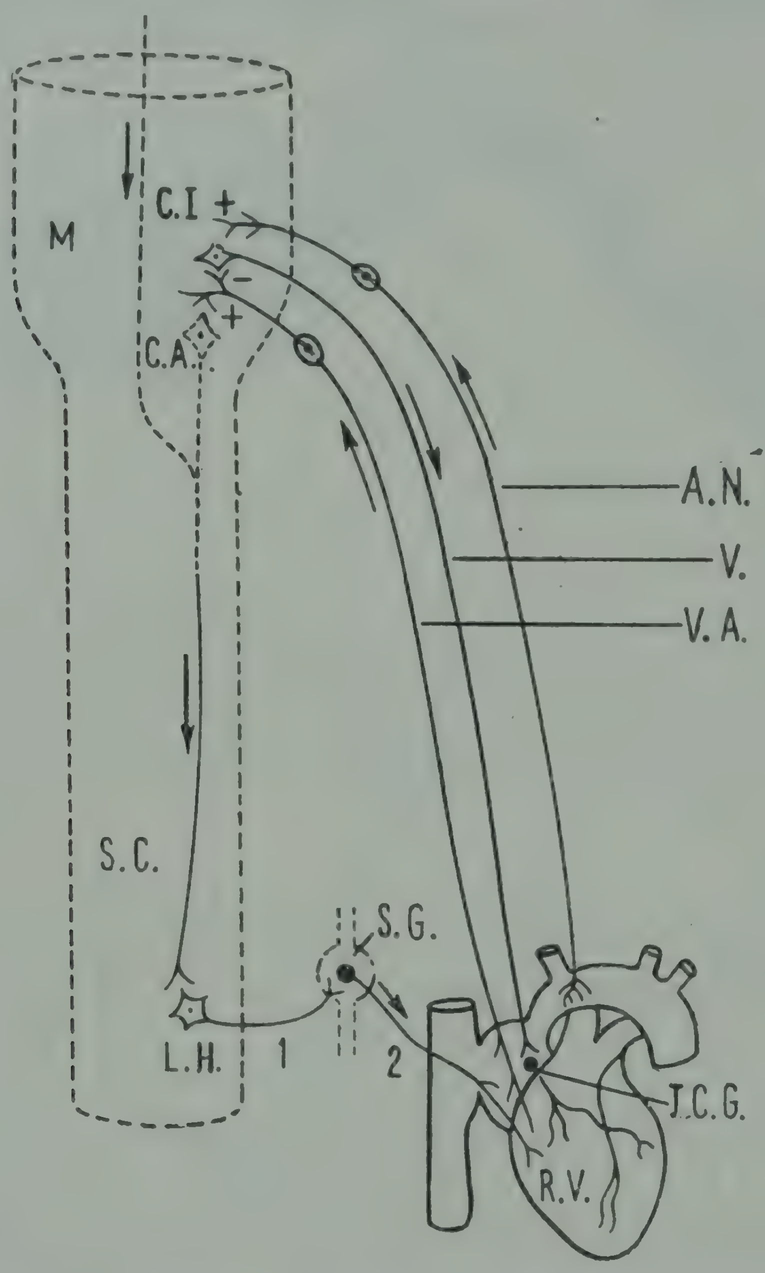


Fig. 385. Diagram of the Innervation of the Heart.

M. Medulla; S.C. Spinal cord; C.I. Cardio-inhibitory centre; C.A. Cardio-accelerator centre; L.H. Lateral horn cell; S.G. Stellate ganglion; 1, 2, pre- and post-ganglionic fibres; R.V. Right ventricle; V.A. Vagal accelerator afferent; A.N. Aortic nerve; V. Efferent vagal fibre; I.C.G. Intra-cardiac ganglion.

The Efferent Cardiac Nerves. The heart in vertebrates is supplied with efferent nerve fibres from two sources: from the medulla oblongata along the vagus nerve, and from the upper dorsal region of the spinal cord through the mediation of the sympathetic system (Fig. 385).

The Vagus. In the frog, the fibres which leave the vagus to pass to the heart make connection with the cells of Remak's ganglion, and probably with other intrinsic cardiac ganglia, whence non-medullated fibres carry their impulses to the heart muscle. The site of these synapses can be demonstrated by the nicotine method. After painting nicotine on the heart in the region of Remak's ganglion, stimulation of the vagus is without effect on the heart; stimulation of Remak's ganglion or of the fibres leaving it is, however, still effective.

The fibres of the vagus in the mammal arise from the dorsal nucleus of the vagus, and various branches pass to the heart by way of the superficial and deep cardiac plexuses, the anatomical arrangements of which are intricate. In the mammal, the right vagus arborises in the sino-auricular node, and the left nerve in the auriculo-ventricular node. As in the frog, the vagus fibres relay with ganglion cells at the places named, there being numerous scattered ganglia in both auricles and in the auriculo-ventricular groove, but few, if any, in the ventricles.* From these nerve cells, fibres are distributed to auricles, to the bundle of His and to the base of the ventricles.

THE SYMPATHETIC. The fibres from the sympathetic system take a somewhat different course in the frog and the mammal. In the frog (Fig. 386) the sympathetic fibres leave the spinal cord by the anterior root of the third spinal nerve; they then pass through the ramus communicans to the corresponding sympathetic ganglion, whence they run up through the second ganglion and the annulus of Vieussens to the first ganglion; they then pass into the cervical sympathetic strand to the *ganglion trunci vagi;* here they join the vagus and pass down as a vago-sympathetic to the heart.

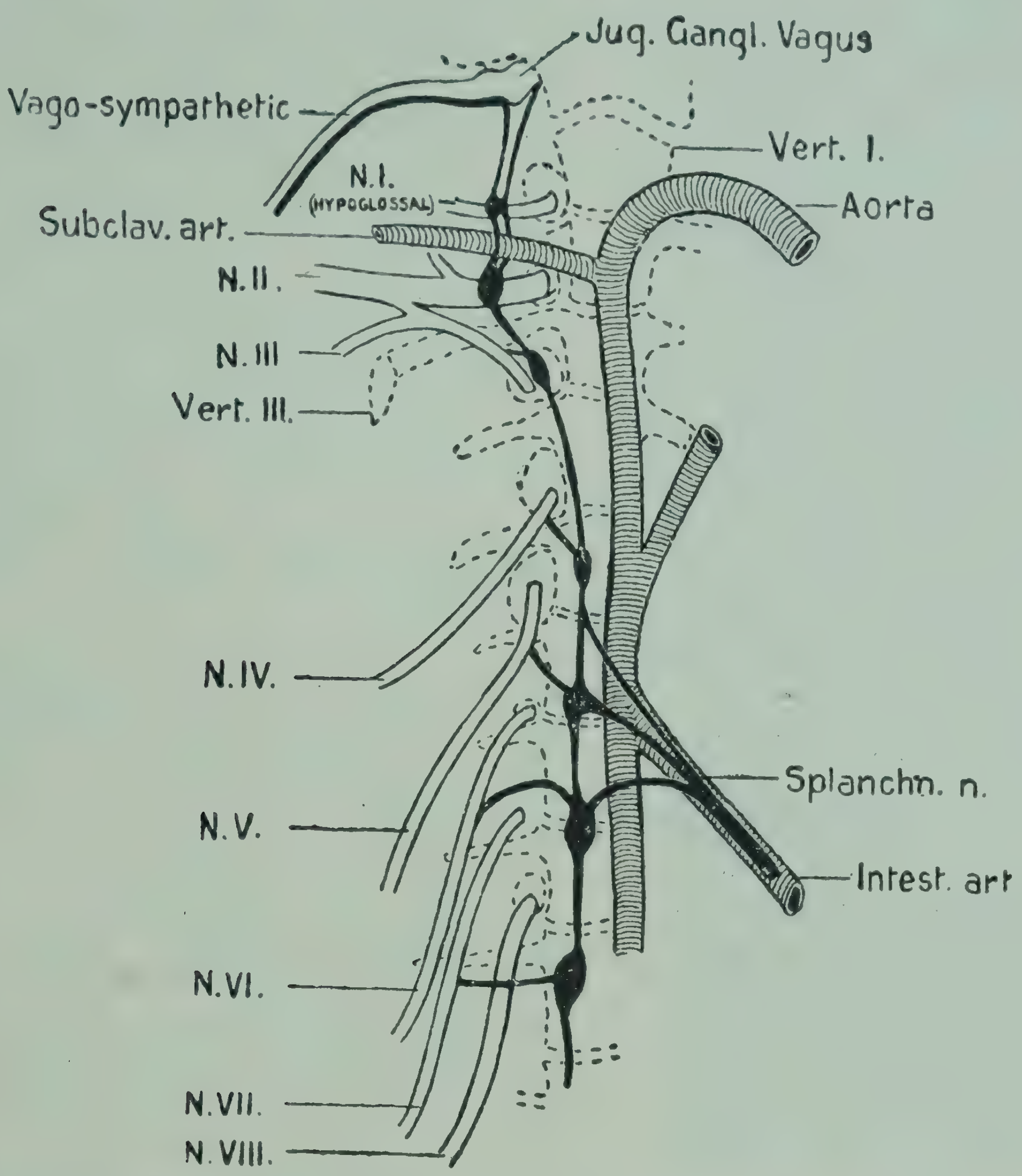


FIG. 386. Sympathetic Chain of Frog (right side) to show Connection with Vagus Nerve.

The sympathetic ganglia with their branches are black. Of the peripheral branches only the splanchnic nerve is represented. (Modified from ECKER.)

In most mammals the principal sympathetic fibres leave the spinal cord by the ventral roots of the upper 5, but chiefly the second and third thoracic nerves, run in white rami communicantes to the stellate ganglion or its equivalent (= inf. cerv. g. + 1–4 thoracic g.), and thence by the annulus of Vieussens to the middle cervical ganglion. Cardiac branches convey the sympathetic fibres to the heart, and are given off from the stellate ganglion, the middle cervical ganglion and the annulus. These branches run to the superficial and deep cardiac plexuses, in which they cannot be distinguished from the vagus branches.

By the nicotine method it is possible to trace out the cell connections of these fibres. As they leave the cord they are medullated, and similar to the other fibres making up the visceral outflow throughout the thoracic region; the white fibres pass along the ramus communicans to the stellate or other sympathetic ganglion, where they end, forming *synapses* with the cells of the ganglion. Here fresh relays of fibres, which are non-medullated, start and carry the impulses to the heart along the various cardiac nerves just mentioned. In the heart these fibres are distributed to the muscle fibres without the intervention of any other ganglion cells.

Although the principal supply of sympathetic fibres to the heart is as stated above, there are subsidiary supplies derived from the white rami of the thoracic spinal nerves down to the 6th or 7th, while cardiac branches from the sympathetic chain run from the superior, middle and inferior cervical ganglia as well as from the upper thoracic ganglia, to the heart. This probably applies also to man (Fig. 387), though it is not certain that all these branches are cardio-accelerator in function.

---

* WOOLLARD. *J. Anat.*, 1926, **60**, 345.

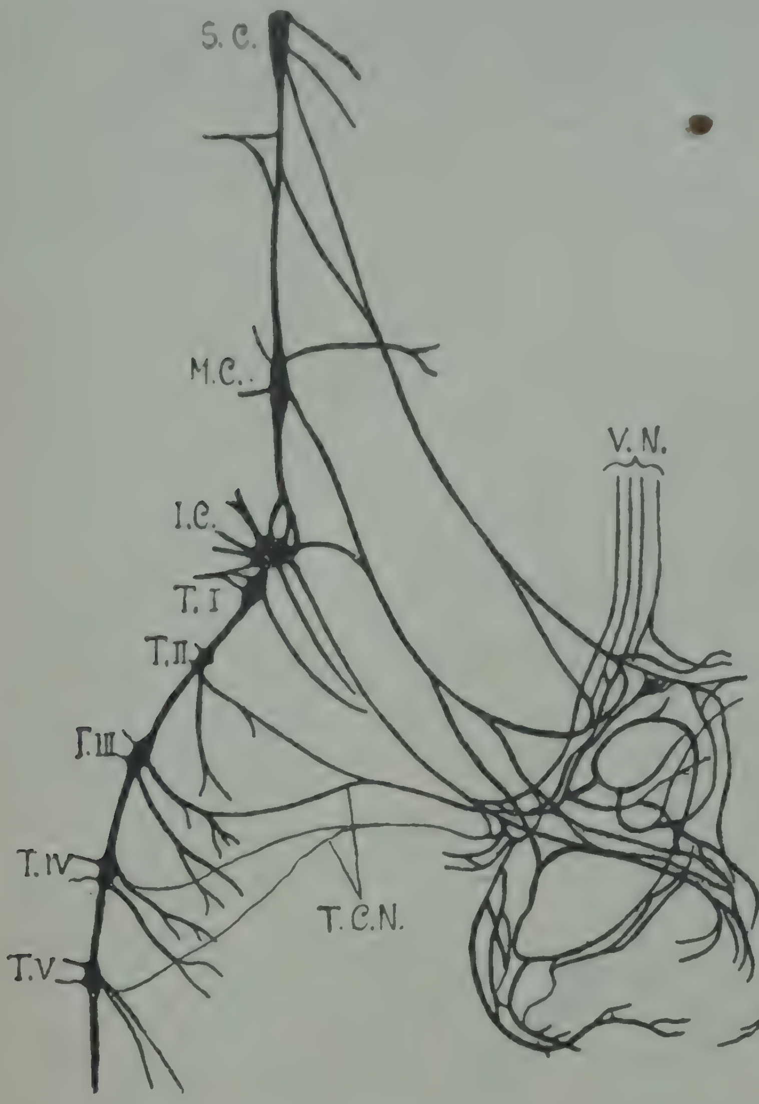


FIG. 387. Diagram of Sympathetic and Vagus Supply to the Cardiac Plexuses (Man). S.C., M.C., I.C., superior, middle and inferior cervical ganglia. T. I, II, III, IV, V, thoracic sympathetic ganglia. T.C.N., thoracic cardiac nerves. V.N., cardiac branches of vagus. (KUNTZ.)

Action of the Vagus. The action of the vagus fibres on the heart is almost identical in frog and mammal and is exerted primarily on the auricle. Since the auricle normally drives the ventricle, this also is affected. If in the dog the peripheral end of the cut vagus be stimulated, while the arterial blood pressure is being recorded, diastole is lengthened almost immediately, the pulse is seen to become slower, or, with a stronger stimulus, to cease altogether, and the blood pressure falls towards zero. On discontinuing the stimulus, the heart begins to beat again and the pressure rises after a few beats to normal (Fig. 388).

If the stimulation of the vagus be prolonged, the blood pressure, on discontinuance of the stimulus, may rise above normal owing to the asphyxia of the vaso-motor centres produced by the prolonged cessation of the circulation.

When the arterial system is dilated, so that the mean systemic pressure during cardiac inhibition is low, or when the asphyxial gasps of the animal are prevented by section of the spinal cord, the heart may fail to recover from the inhibition produced by stimulation of the vagus.

To study the influence of the vagus on the auricles and ventricles respectively, it is necessary to record separately the contractions of the different chambers of the heart, or the electrocardiogram may be used as an index to the changes in the heart under the action of nerve stimulation or section. The vagi may affect the heart in one of several ways. Their chief action is on the venous end. Vagus stimulation may cause auricular standstill, slowing, flutter or fibrillation. The right vagus, since it connects with the sino-auricular node, has the greater effect in lowering the rate. If its action is confined to the sino-auricular node, the sole effect of the vagus on the auricles and ventricles will consist in an alteration of rhythm. They may cease to beat altogether, or they may give beats of normal strength but at a slower rhythm than before. Often, indeed, under these conditions the beats of the ventricles may be increased in size, since the strength and extent of their contractions will be greater with any prolongation of the diastolic period and consequent increased diastolic filling of the ventricle.

If the vagus (especially the left) acts on the auricles without affecting the sino-auricular node, the rhythm will be unaltered, but the strength and duration of systole, and the refractory period of the auricles, will be diminished. Under such circumstances the rhythm of the ventricles will be unchanged.

Generally, strong vagus stimulation stops the action of all parts of the auricles ; in such cases the ventricles also cease beating. Very often after a short pause the ventricles commence to beat at a slow (" idioventricular ") rhythm. This is called " vagus escape." This escape is generally confined to the ventricles, the auricles remaining in a state of diastole. Vagus escape is favoured by distension of the right side of the heart, and is often synchronous with the respiratory efforts which supervene after a certain duration of inhibition, as a result of asphyxia. It appears to be due in part to an accelerating reflex from a distended right auricle,* and from the carotid sinus in which pressure has fallen, and in part to independent contraction of a ventricle temporarily cut off from auricular control. That the ventricle can initiate the escape beat is shown by the fact that occasionally one may observe a contraction of the auricle following each ventricular contraction.

Whether the vagus has a direct action on the mammalian ventricle is still doubtful; its effect is, at any rate, very slight as compared with that on the auricle. The fact that stimulation of the vagus, like division of the bundle of His, usually causes merely temporary cessation of the ventricular beat, also indicates that this nerve has its chief action on the auricles.†

In the frog, it appears that the vagus has a direct inhibitory action on the ventricle, diminishing the strength of its contraction in response to the stimuli transmitted to it from the venous end. This action of the vagus on the ventricle is, however, not found in the tortoise. In both these animals the auricles show the same effects as in the mammal, viz. an influence limited to the rhythm when only the sinus is affected, or a diminution of the strength of contraction when the sinus is unaffected and when the chief action is on the auricular muscle.

Finally, the vagi may affect the conduction of the excitatory process from auricles to ventricles and to all parts of the ventricles.‡ Under weak vagus stimulation this is shown by increase in the P–R interval of the electrocardiogram, and in more pronounced effects the ventricles may beat at a slower rhythm than the auricles, a block having been produced in the auriculo-ventricular bundle.

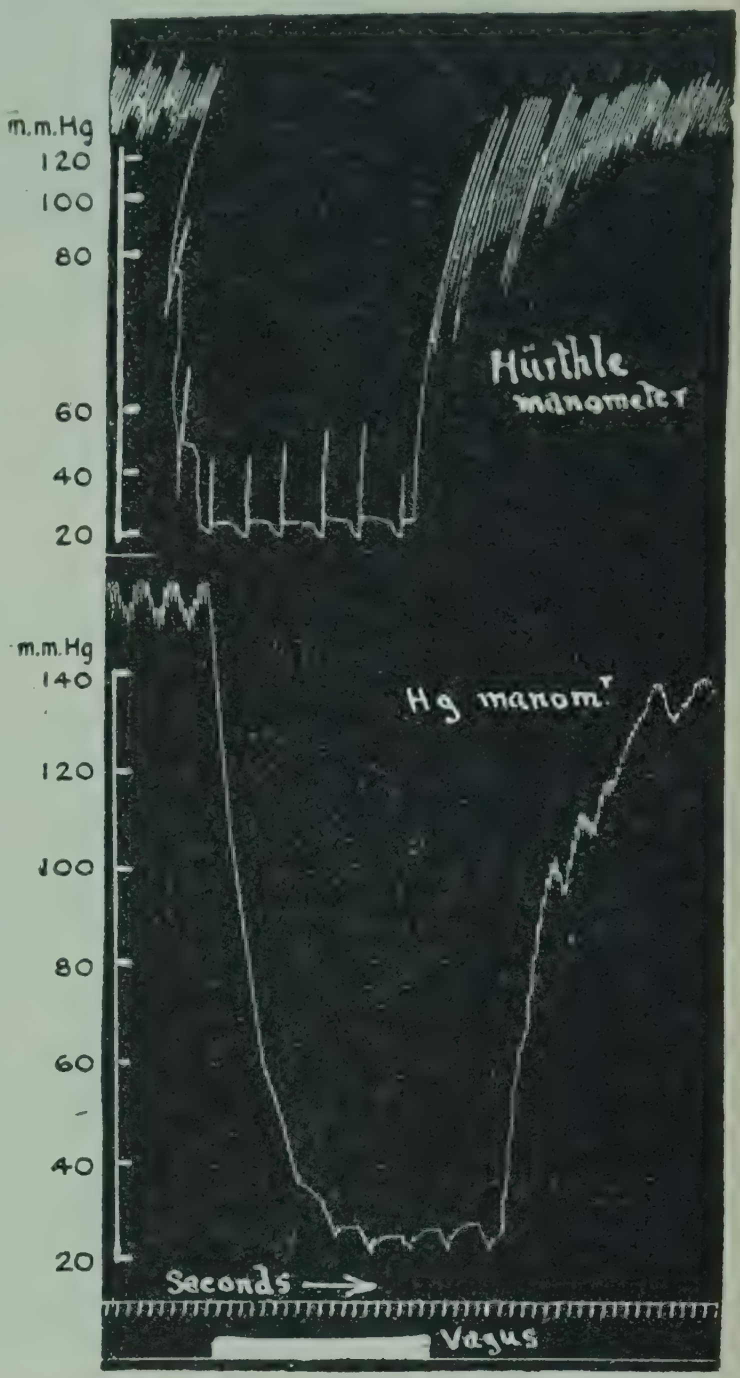


FIG. 388. Blood Pressure Tracing from Carotid of Dog (taken simultaneously with Hürthle's Manometer and Mercury Manometer), showing effect of Excitation of Vagus. (The Mercury Manometer gives very imperfect records of the pulsations during rapid changes of blood pressure.) (C. L. E.)

It is evident that these various effects are chiefly due to differences in the place affected. Thus, if the vagus fibres which are distributed to the remains of the sinus are specially active, we shall get alterations of rhythm affecting the whole heart. If those which supply the A.V. bundle are excited, the most pronounced effect will be on the propagation of the excitatory process from auricles to ventricles.

It is remarkable that the excitability of the vagus fibres is dependent on the condition of the cells in the medullary centres. Thus, if the centres are depressed by narcotics, the vagus fibres become more excitable, and similarly if the vagi are sectioned, so that the fibres are cut off from the centres, their excitability steadily increases to a maximum.§

The vagus nerve usually contains also some cardio-accelerator fibres, which arise directly from the medulla.||

THE TONIC ACTION OF THE VAGUS. If both vagi of a mammal be divided, or if a small dose of atropine be given, the heart as a rule beats more frequently,

---

* McDOWALL. *J. Physiol.*, 1926, **61**, 131.

† ERLANGER and HIRSCHFELDER. *Amer. J. Physiol.*, 1906, **15**, 159; CULLIS and TRIBE, *J. Physiol.*, 1913, **46**, 141.

‡ COHN. *J. Exp. Med.*, 1912, **16**, 732; COHN and LEWIS, *J. Exp. Med.*, 1913, **18**, 739.

§ GREMELS. *Arch. f. exp. Path.*, 1935, **179**, 360.

|| PANNIER. *Arch. int. pharmacodyn.*, 1946, **73**, 193.

showing that, under normal circumstances, tonic impulses are constantly descending the vagi and holding the heart's action in check. The extent of the quickening produced by section of the vagi varies in different species and is associated with its powers of carrying out prolonged muscular exertion. Thus, in the dog, hare, or horse the pulse, which is normally slow, may be doubled in frequency by section of the vagi. In the rabbit, which has a rapid pulse, division of both vagi causes very little alteration in the pulse rate. In man, especially in the athlete, vagus tone is high.

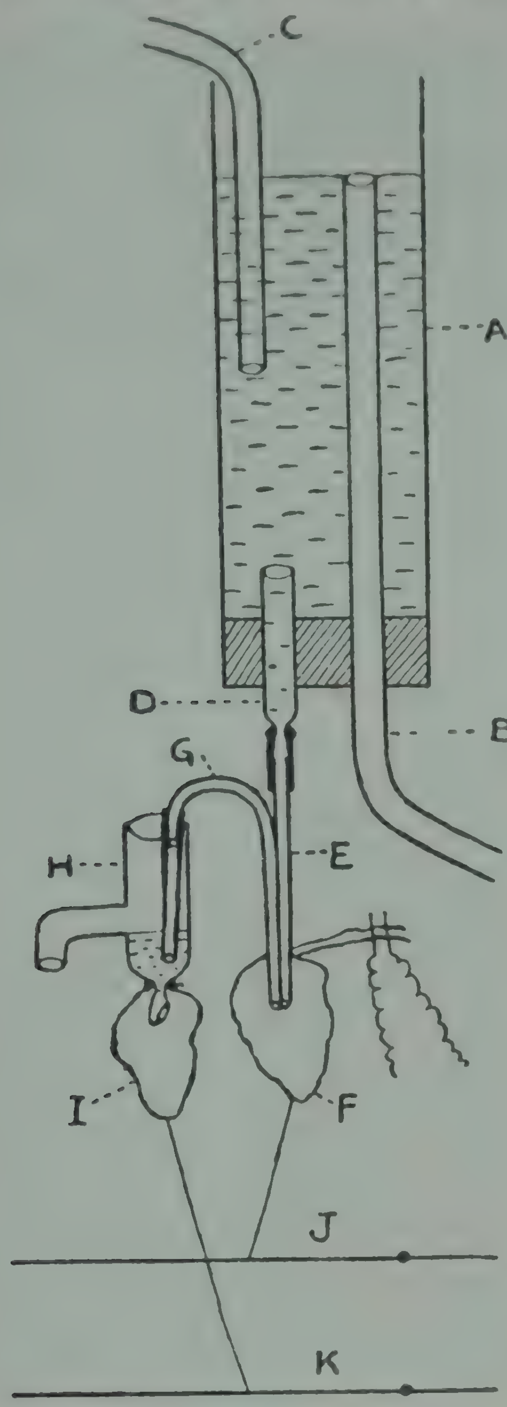


FIG. 389. The fluid from perfusion apparatus A passes to E which supplies the donor heart F. After irrigating the inside of the donor heart, the fluid passes to the glass cannula tube H, to which is attached the recipient heart I. J and K are the levers to which the hearts are attached. (BAIN. *Q. J. Exp. Physiol.*, 1932, **22**, 269.)

This tonic action may be increased by morphia. Reflex alteration of the tone may also be induced in various ways to be described later, and is normally the chief cause of alteration in heart rate.

MODE OF ACTION OF VAGUS. Ever since the discovery in 1845 by the brothers E. H. and E. F. Weber * of the action of the vagus on the heart, much work has been expended with a view to determining the intimate nature of the inhibitory process. Atropine is known to paralyse the effect of stimulation of either the vagus trunk or the ganglion of Remak in the frog. It therefore exerts its action at or beyond the post-ganglionic terminations. Muscarine or acetyl choline produces the same effect as stimulation of the inhibitory fibres, and when applied to the isolated auricle or ventricle causes weakening of the beat and finally complete inhibition; this action is also prevented by atropine.

Although various views as to the nature of inhibition have been put forward from time to time, the present tendency is to ascribe both the excitatory and the inhibitory effects of all nervous stimulation to the liberation, at the nerve ending, of small amounts of highly active substances, which exert their effects on, and are quickly destroyed in, the tissues.

It was shown by Loewi † that the vagus acts by producing in the heart a chemical substance which has the power of inhibiting it. The "humoral" transmission of the effect of vagus stimulation can be demonstrated as in Fig. 389. When a frog's heart perfused with saline solution is brought to a standstill by stimulation of its vagus, the fluid if passed on into a fresh heart will produce on the latter the same effect as if the vagus nerve to the second heart were stimulated (Fig. 390). It is now generally accepted that the mediating substance is acetyl choline, and that the effect of vagus excitation is to produce or liberate this substance in the immediate vicinity of the vagus endings.

---

* E. H. WEBER. *Annali. universali di medicina*, 1845, **116**, 225; E. F. WEBER in Wagner's *Handwörterbuch d. Physiol.*, 1846, **3** (ii), 42.

† LOEWI. *Pflüger's Arch.*, 1921, **189**, 239.

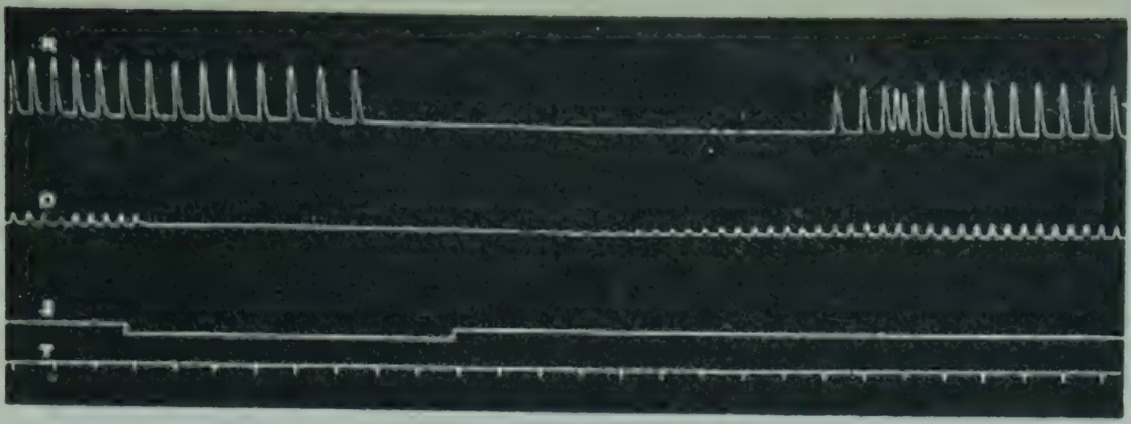


FIG. 390. Tracings from hearts of *Rana esculenta* perfused by method illustrated in Fig. 389. Upper line, recipient heart (R); second line, donor (nerve-stimulated) heart (D); third line, signal; bottom line, time in five-second intervals. Vagus of donor heart stimulated during forty seconds. Arrest of donor heart followed by slowing and then by standstill of recipient heart. (BAIN. *Q. J. Exp. Physiol.*, 1932, **22**, 269.)

The action of acetyl choline is antagonised by atropine, but augmented and prolonged by eserine. The latter effect is due to the fact that eserine paralyses the activity of the enzyme choline esterase, which is prevalent in the blood and tissues, and which normally very rapidly destroys the acetyl choline, so shortening the time over which its effects will last.

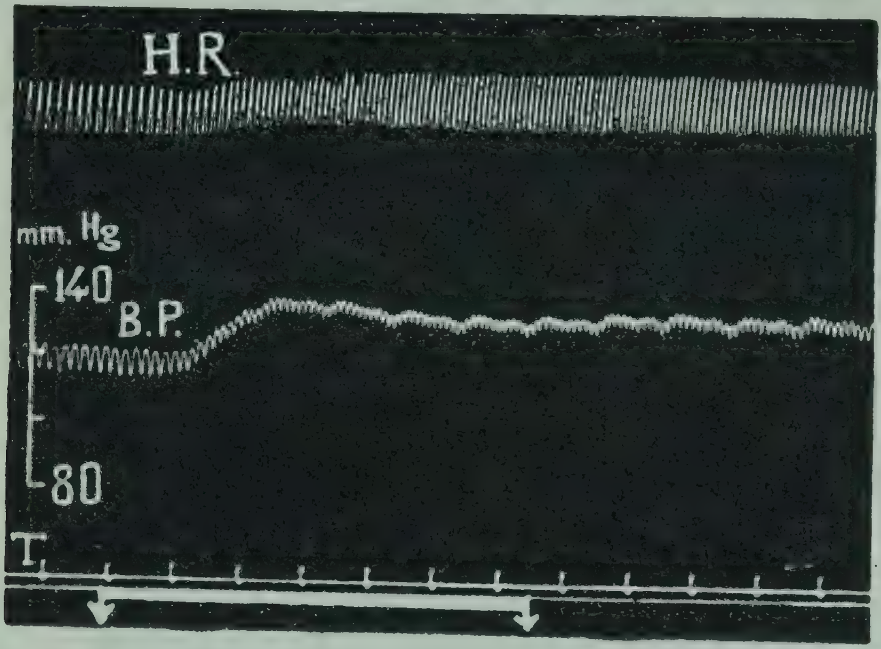


FIG. 391. Heart rate (H.R.) as recorded by Hürthle manometer, and arterial pressure (B.P.) of decerebrate dog, with suprarenals removed and all cardiac nerves cut. Between the arrows stimulation of cardiac nerves from right stellate ganglion causes acceleration of 93 beats per minute and rise of arterial pressure. T = 3-second intervals. Note the latent period of about 6 seconds. (SAMAAN. *J. Physiol.*, 1935, **83**, 332.)

Action of the Sympathetic. Stimulation of the sympathetic cardiac nerves at any part of their course has an effect on the heart the reverse of that produced by stimulating the vagi.* In most cases the pulse frequency is increased in consequence of the action of these nerves on the "pace-maker" of the heart. The frequency attained by maximal stimulation of the accelerator nerves is independent of the previous rate of the heart beat. The increase in rate involves a shortening of the diastolic period. The size of the auricular and ventricular contractions may be increased at the same time as their rate. In fact, like the vagus nerves, the sympathetic fibres of the heart can influence rhythm, strength of contraction, or conduction from auricle to ventricle, according to the part of the heart muscle which is affected.

The augmentor effect on the strength of the ventricular beats is often great. The sympathetic fibres are much less easily tired than the vagus fibres, and have a longer latent period. Whereas the latent period of the vagus in the mammal is considerably less than one second, that of the accelerator nerves may amount to ten seconds (Fig. 391). If the vago-sympathetic of the frog be stimulated, the first effect is inhibition due to vagus action. The vagus nerve endings then become fatigued, and the influence of the accelerator fibres makes

---

* CYON and CYON. *Arch. f. Anat. u. Physiol.*, 1867, 389.

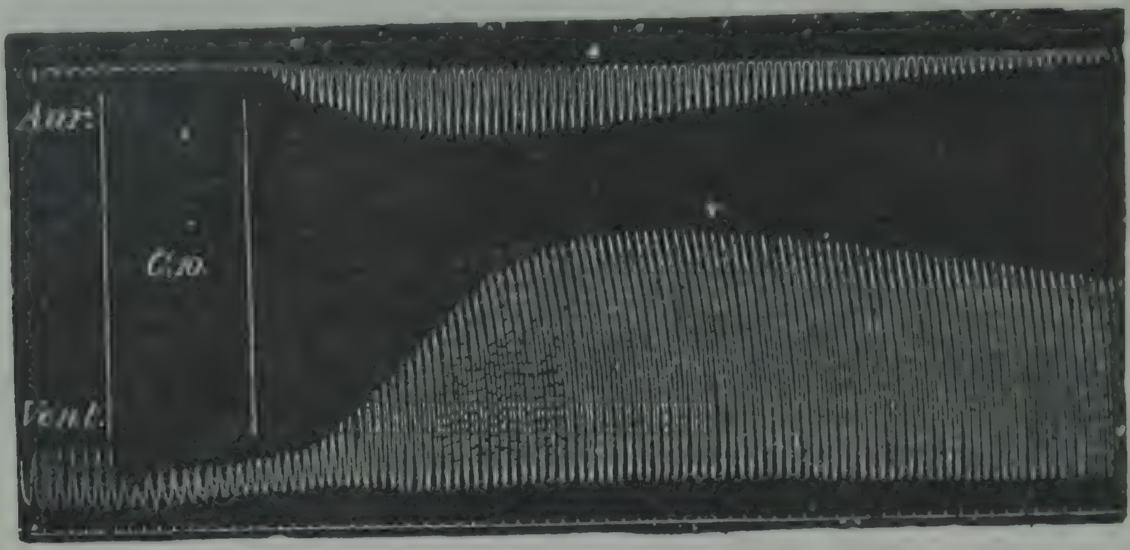


FIG. 392. Tracing to show Effect of Stimulation of the Vago-Sympathetic Nerve on the Frog's Heart. (GASKELL.)

The rhythm is unaltered, but the beats of auricle and ventricles are much decreased in size. On ceasing the stimulation the beats become augmented.

itself apparent ; the heart commences to beat, and the beats become more rapid and forcible than before * especially just after stimulation ends (Fig. 392).

Like the vagus, the sympathetic nerve fibres appear able to exercise a tonic influence on the heart so that, after extirpation of the stellate ganglion on each side, the pulse may become permanently, though only slightly, slowed.

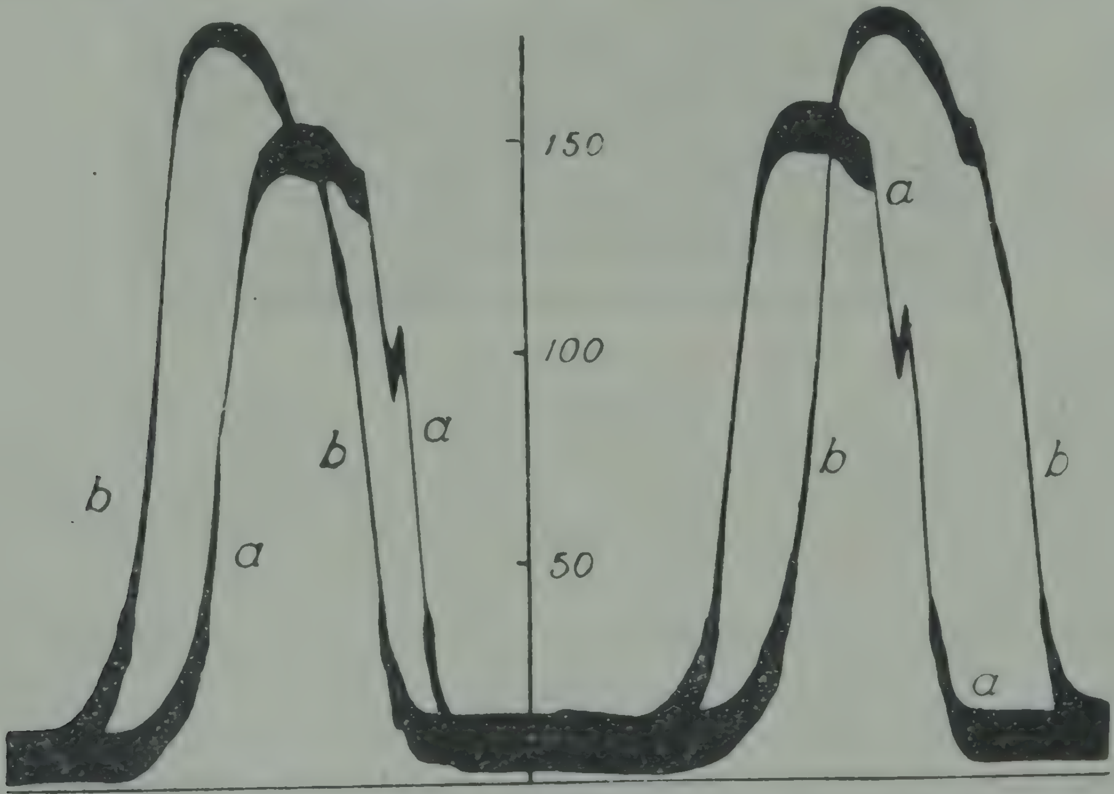


FIG. 393. Intraventricular Pressure Tracings (Left Ventricle) from Dog's Heart (Heart-lung Preparation). (S. W. PATTERSON.)

(To be read from right to left.) The scale shows pressure in mm. Hg.
a. Under influence of adrenaline.
b. Under simultaneous influence of adrenaline and CO_2 (15 per cent.).

THE ACTION OF ADRENALINE ON THE HEART. Adrenaline has a powerful action both on the heart and blood vessels and plays an important part in the regulation of the circulation. The action is practically identical with that of excitation of the sympathetic nerve supply. Indeed, as a parallel to the functions of acetyl choline in the parasympathetic system, it is probable that the action of the sympathetic nerves is actually due to the liberation of small amounts of adrenaline at the nerve endings, so that a local effect is produced. Its action on the heart is well shown in the heart-lung preparation. On adding $\frac{1}{10}$ mgm. of adrenaline to the 500 c.c. of blood circulating through the preparation, a maximum accelerator and augmentor effect is produced and lasts for 15 to 20 minutes. Through its influence on the sino-auricular node, the rhythm of the heart is greatly increased, in the dog to about 240 per minute. At the same time the energy of each contraction is augmented. This is especially shown in a heart which is beginning to fail and is therefore showing dilatation. Directly the adrenaline reaches the heart, the contractions become extremely energetic so that the heart rapidly diminishes in volume, the venous pressure falls, and the blood is thrown out with violence into the aorta. The more powerful contraction enables the output of the heart at each beat to be maintained or even increased, in spite of the shorter duration of the systole. With each beat the maximum pressure in the ventricle therefore rises (*see* Fig. 393). The stimulating effect of adrenaline is shown, moreover, by the fact that the oxygen usage is increased two or three times above that which obtained before the administration * ; it also dilates the coronary vessels, so ensuring an increased supply of blood to the heart. The action of adrenaline, therefore enables the heart to cope with a bigger strain, either in the shape of arterial resistance or increased inflow, than it could otherwise do.

---

* GASKELL. *J. Physiol.*, 1884, **5**, 46 ; Proc. ix.

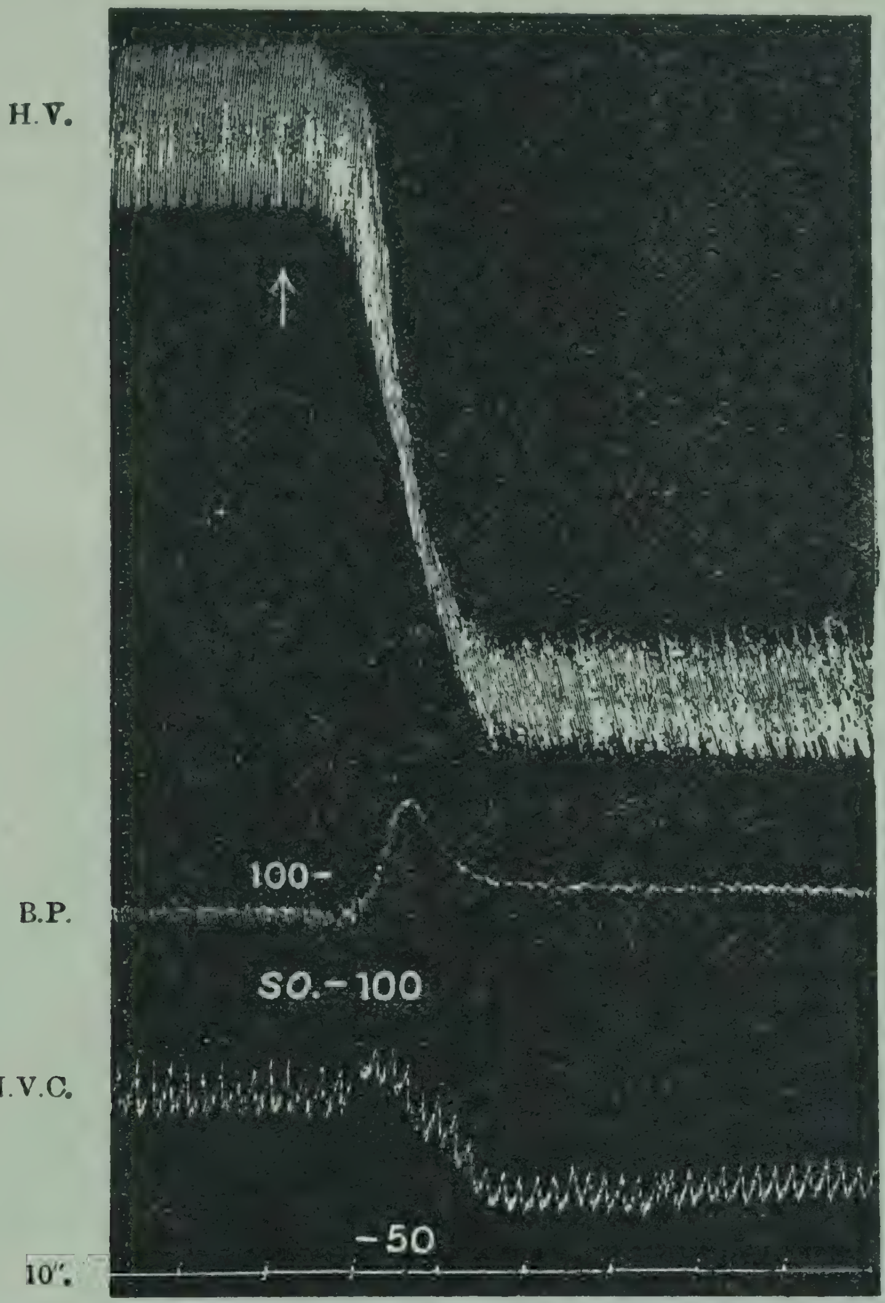


FIG. 394. Effect of Adrenaline on Heart Volume and Venous Pressure. Dog Heart-lung preparation. H.V. = heart volume ; B.P. = arterial pressure ; I.V.C. = venous pressure. At arrow adrenaline added to give a concentration of 1 : 3,500,000 in the blood. (PATTERSON, *Proc. R. S.*, 1914, **88B,** 371.)

There is an antagonism in the influence on the heart of the two substances set free during general muscular activity. Carbon dioxide in excess causes diminished activity, dilatation of the heart and slowing of rhythm. Adrenaline causes increased activity, diminution of cardiac volume and increased rhythm (Fig. 394). The action is so pronounced that if adrenaline be added at the same time,* it is possible to administer 20 or 30 per cent. carbon dioxide to a heart-lung preparation without altering its output, and the beat is maintained and contraction is effective.

---

* EVANS and OGAWA. *J. Physiol.*, 1914, **47**, 446.

THE HEART REFLEXES

Centres. These are in the medulla oblongata. Here, in the dorsal nucleus of the vagus, we find the *cardio-inhibitory centre, i.e.* the neurone pool giving origin to the efferent fibres of the vagus, and also in which the afferent fibres of the vagus terminate.† Direct stimulation of the vagus centre may cause slowing or stoppage of the heart. The tonic influence of the vagi is abolished by destruction of this centre. In this region we also find the vasomotor centre, and the two centres are often affected together, though in opposite directions, excitation of the cardio-inhibitory centre for instance being accompanied by a reduced activity of the vasomotor centre. The cardio-inhibitory centre may be played upon by impulses, arriving at it through various afferent nerves or from the higher parts of the brain, which give rise to the changes of the pulse rate associated with emotional conditions, or it may be directly affected by the composition of the blood circulating through its capillaries.

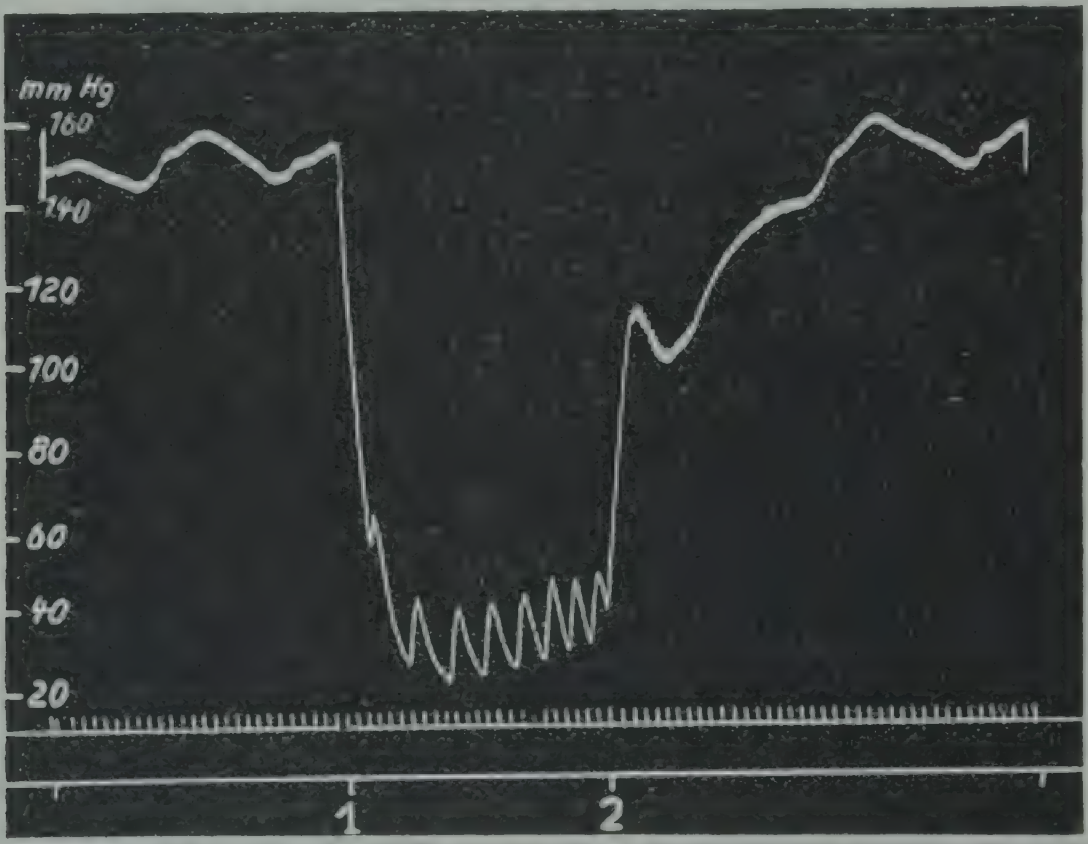


FIG. 395. Blood pressure of Dog. Stimulation of left aortic nerve between 1 and 2. Time = seconds. (SCHWEITZER. Die Irradiation autonomer Reflexe, 1937. Basel, Karger).

The nerve cells which give off the sympathetic accelerator fibres are situated in the intermedio-lateral tract of the spinal cord, near the roots of origin of these fibres. We might, therefore, speak of a *cardio-accelerator* centre in this region ; but it is most probable that the activity of these cells is subordinated to impulses arriving at them from a common site for control of visceral impulses, viz. the medulla, or even higher, as in the hypothalamic region.

Most sensory nerves of the body when stimulated give either a slowing or a quickening of the heart, *e.g.*, stimulation of the fifth nerve, as in the nasal mucous membrane, sometimes causes reflex inhibition.

In the various reflexes in which these cardiac centres are involved, this happens in a reciprocal manner, *e.g.* increase of action of the accelerator centre being accompanied by a reduced activity of the cardio-inhibitory centre.

---

* PATTERSON. *Proc. Roy. Soc.*, 1914, **88B**, 371.

† MILLER and BOWMAN. *Amer. J. Physiol.*, 1915, **39**, 149.

The most important afferent nerves which affect the heart reflexly are the cardiac afferents, from the aorta and heart itself, and the carotid sinus nerve.

Cardiac Afferents. The heart does not seem to be provided with nerves of ordinary sensibility. There is no doubt, however, that under abnormal circumstances, impulses arising in the heart can give rise to sensations of pain, which are referred not so much to the heart as to the surface of the body over the left side of the chest and left arm, in the region of the distribution of the cutaneous branches of the second and third dorsal roots. The nervous impulses responsible for these anginal pains have been said to be conveyed by sympathetic afferents.

THE AORTIC NERVE. The chief afferent nerve coming from the beginning of the aorta, is the *cardio-aortic*, or *depressor* nerve.[*] Neither name is quite satisfactory. It is usually bound up with the vagus, but many of its fibres form a separate nerve in the rabbit. This connects with the superior laryngeal branch of the vagus and with the upper vagus trunk, and runs parallel with the vagus to the cardiac plexus. In man it is represented by a twig of the external branch of the superior laryngeal nerve.

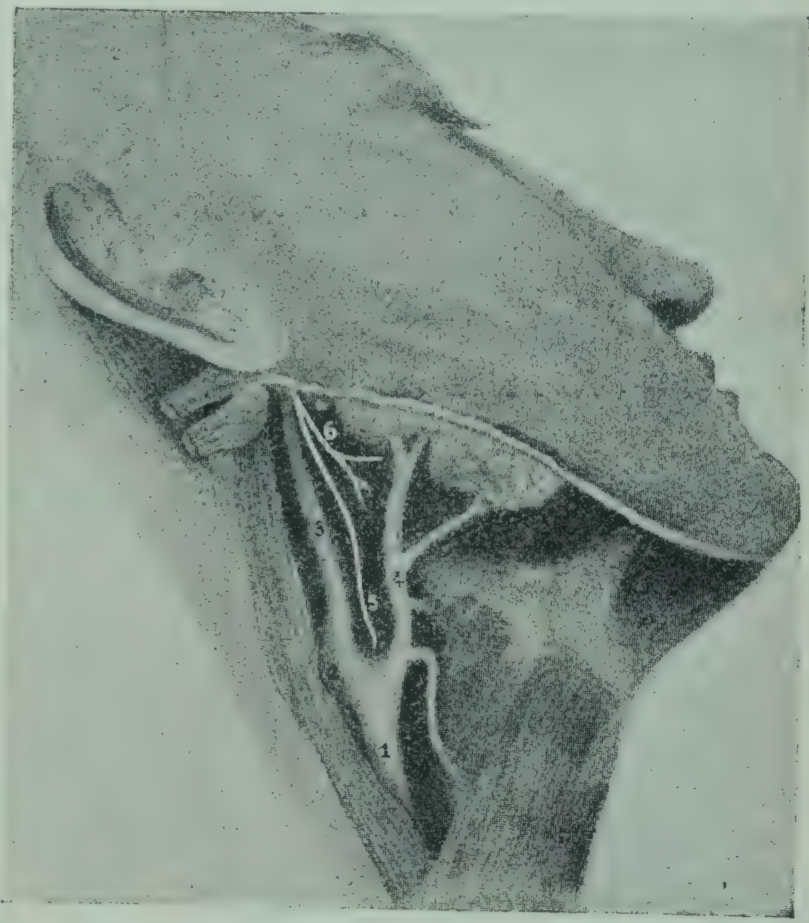


FIG. 396. Showing carotid sinus in man: (1) common carotid; (2) carotid sinus; (3) internal carotid; (4) external carotid; (5) nerve to carotid sinus; (6) glosso-pharyngeal nerve. (C. HEYMANS, "Le sinus carotidien.")

The cardio-aortic nerve arises from endings in the aortic arch and in the left ventricle. Similar afferents originate in the pulmonary artery, and have similar, but weaker, functions[†]; fibres in the pulmonary veins, also probably vagal, are more potent.[‡] These endings are sensitive to stretch.

The aortic nerve is purely afferent, stimulation of its peripheral end causing no effect. On stimulating its central end, or the central end of one vagus, preferably the left, a fall of arterial blood pressure (Fig. 395) and reflex slowing of the heart are produced, the latter effect being abolished by section of both vagi. The reflex slowing is due to impulses which ascend by the cardio-aortic fibres of the vagus and are then returned from the vagus centre *viâ* the cardio-inhibitory fibres of the vagus. The fall of blood pressure still occurs on central vagal excitation, when both vagi are divided. It was shown by Bayliss that[§] this depressor effect is due to widespread dilatation of the blood vessels of the body, chiefly in the splanchnic area.

The aortic nerve is brought into action by stretching or dilating the aorta, as happens when the blood pressure is high. At each systole a rapid shower of nerve impulses can be shown to pass up the afferent fibres of the vagus or aortic nerve; when the aortic pressure is higher these impulses are more numerous. Synchronously with these, a shower of outward impulses passes down the efferent fibres of the vagus. If the depressor nerves be cut when the aortic pressure is high, the blood pressure rises still higher. We have here a means by which the heart can be relieved of a load too great for its powers.

---

[*] LUDWIG and CYON. *Ber. Sächs. Ges.*, 1866, 307.

[†] SCHWEITZER. *J. Physiol.*, 1936, **87**, 46P.

[‡] DALY *et al.* *Q. J. Exp. Physiol.*, 1937, **27**, 123.

[§] BAYLISS. *J. Physiol.*, 1893, **14**, 303; 1908, **37**, 264; *Proc. Roy. Soc.*, 1908, **80B**, 339.

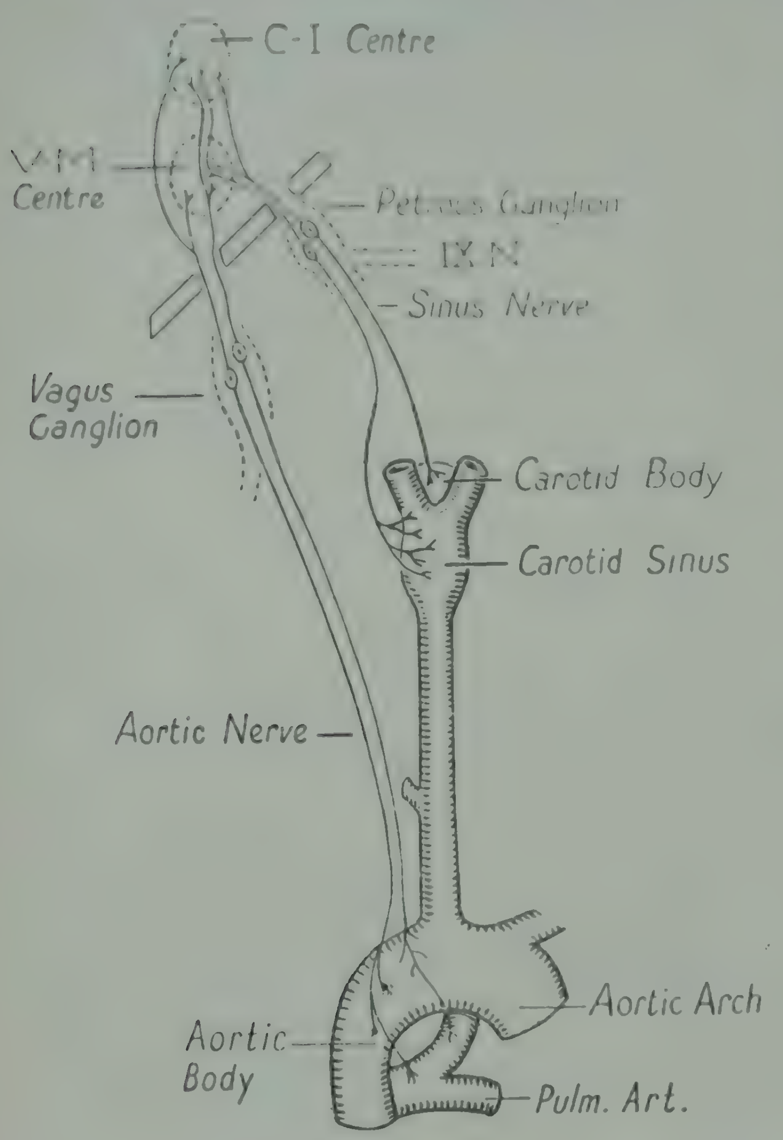


FIG. 397. Diagram of the afferent innervation of the carotid and aortic regions in relation to cardio-vascular reflexes.

The *aortic bodies*, patches of vascular tissue on the aortes, and similar in structure and function to the carotid body, also send fibres to the cardio-aortic nerve. The nerve endings from which these fibres arise are sensitive to chemical changes in the blood, but impulses along them do not affect the heart rate.

OTHER VAGAL AFFERENTS. In the mammalian heart, nerve strands can be seen running over the surface of the ventricle; these are entirely afferent, stimulation of their peripheral ends causing no effect on the heart beat. Stimulation of their central ends may cause one of five results:

(*a*) Slowing of the heart.

(*b*) Acceleration of the heart.

(*c*) Rise of blood pressure from constriction of the splanchnic area.

(*d*) Fall of blood pressure by dilatation of the arterioles of the body.

(*e*) Reflex movements.

These strands ultimately pass up in the vagus nerves.

Some strands of the cardiac plexus cause acceleration of the heart when centrally stimulated. It is probable that these strands originate in the right side of the heart; they reflexly lower vagus tone. Afferent impressions reaching the centres from the lungs *viâ* the vagi also affect the rate of the heart.* Slight inflation of the lungs or weak stimulation of the pulmonary branches of the vagi cause reflex acceleration †; great distension, or strong stimulation causes reflex slowing (*v.* also p. 600).

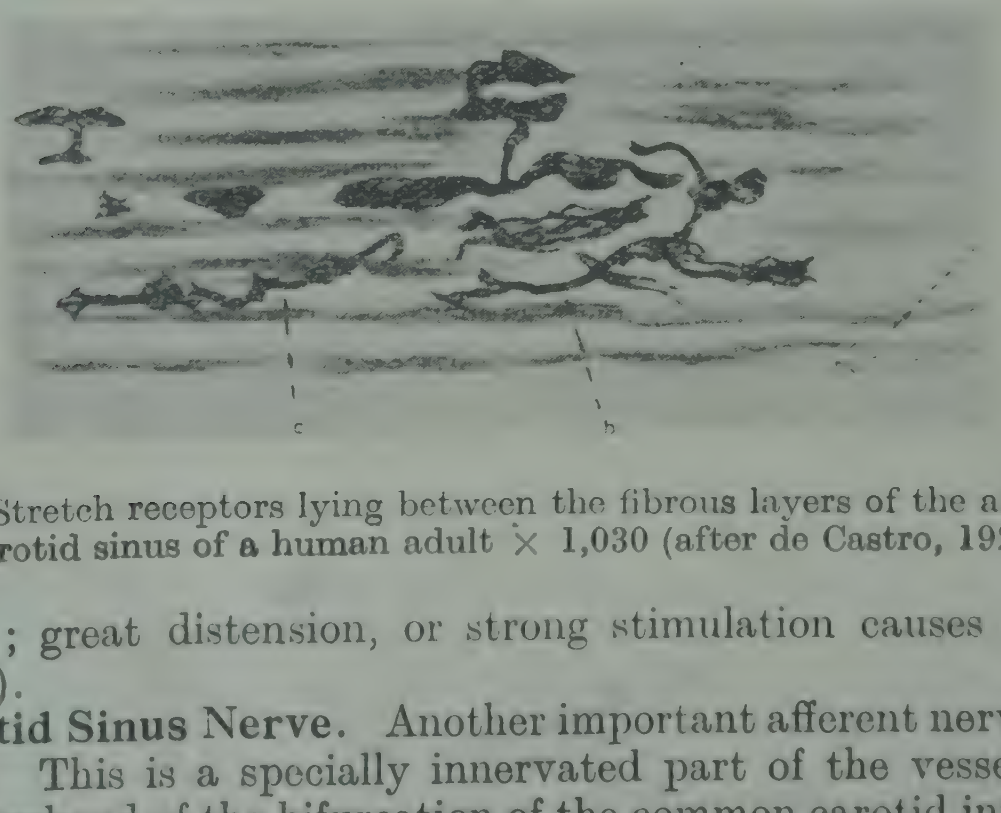


FIG. 398. Stretch receptors lying between the fibrous layers of the adventitia in the carotid sinus of a human adult × 1,030 (after de Castro, 1928).

The Carotid Sinus Nerve. Another important afferent nerve arises in the carotid sinus. This is a specially innervated part of the vessels and tissues in the neighbourhood of the bifurcation of the common carotid into its branches and includes also the carotid body. In man (see Fig. 396), and many other mammals, the common carotid artery divides into the internal and external carotids, and a bulbous thin-walled expansion, the carotid sinus or carotid bulb, is seen at the root of the internal branch. A nerve twig runs from the carotid sinus to the glossopharyngeal (Fig. 397).

---

* SAALFELD. *Pflüger's Arch.*, 1932, **231**, 33.

† ANREP *et al.* *Proc. Roy. Soc.*, 1936, **119B**, 191.

The nerves originating in the carotid sinus are stimulated by stretch, *i.e.* they originate from special stretch receptors (baroceptors) in its walls (Fig. 398).

The carotid body is richly supplied with sinusoids lined with endothelium and peculiar "glomus" cells, the latter being abundantly supplied with nerve fibres that join the carotid sinus nerve (Fig. 397). The nerve endings in the "glomus" cells are not sensitive to stretch, but respond to chemical changes in the blood flowing through the sinusoids.

Impulses set up in the carotid sinus and carotid body produce important reflex changes in heart rate, vascular tone and respiration.

Fig. 397 illustrates diagrammatically the afferent nerve supply responsible for reflex control of the circulation in response to arterial pressure changes, *i.e.* the aortic and carotid sinus innervation. The efferent paths are, of course, the vagus, for cardio-inhibition, and the sympathetic nervous system, connected with the relevant bulbar centres by way of the spinal cord, for cardio-acceleration and, as we shall see later, also for vascular tone.

Stimulation of the carotid sinus electrically, or mechanically, as by stretching its walls by pressure from within in any manner, provokes a combined reflex of cardiac inhibition and fall of blood pressure, just as does stimulation of the central end of the depressor nerve (Figs. 399 and 400). In fact, we may regard the reflexes from the carotid sinus and those from the aortic nerves as part and parcel of the same physiological arrangement for controlling the blood pressure. It is found that, after section of the vagi, stimulation of the carotid sinus in any way no longer produces the cardiac inhibition (proving that this was reflex).

Chemical Responses. Altered chemical content of the blood in the carotid body can produce alterations in heart rate by exciting the chemo-receptors, *e.g.* increase of CO_2 or deficit of oxygen causes reflex quickening of the heart. The effects are relatively weak, however.

Cardiac Reflexes. There are two very important reflex mechanisms associated with the heart itself. If, other conditions remaining the same, the arterial pressure be raised, the heart is slowed. This phenomenon, sometimes dignified with the name of "Marey's law," [*] *i.e.* that *the pulse rate varies inversely as the arterial pressure,* is not invariably seen in the intact animal or in man, because other conditions may be changed. The other reflex is an acceleration of the heart rate when, other conditions remaining the same, the pressure in the great veins and right auricle is raised, as happens when the venous return to the heart is increased. Thus we have two mechanisms, balanced, the one against the other, one causing the heart to slow down if the aorta is getting over full, the other causing it to accelerate as the right side distends.

(i) INHIBITORY REFLEXES. The first of these mechanisms is the reflex cardiac inhibition, resulting when afferent fibres of the aortic nerves (chiefly the left) are excited by the stretching of the arch of the aorta and left ventricular wall if the blood pressure rises. As a result of this the cardio-inhibitory centre is excited, with the sequel that efferent cardio-inhibitory impulses *viâ* the vagi bring about a reduction of the pulse rate.

An equally, or even more, important cause of the cardiac inhibition consequent on rise of arterial pressure is a reflex from the carotid sinus.

This is shown by the experiment called the "perfusion of the carotid sinus." For this purpose, in a dog (B), the sinus region is isolated from the circulation by ligatures, and left connected with the animal only by its nerves; this isolated segment of carotid artery with its branches can then be perfused with blood from a pump, or from a second dog (A). The blood from the carotid of A enters the cephalic end of the carotid artery of B, and, after

---

[*] MAREY. *Physiol. médicale de la circulation du sang,* 1863, p. 206 (Paris, Delahaye).

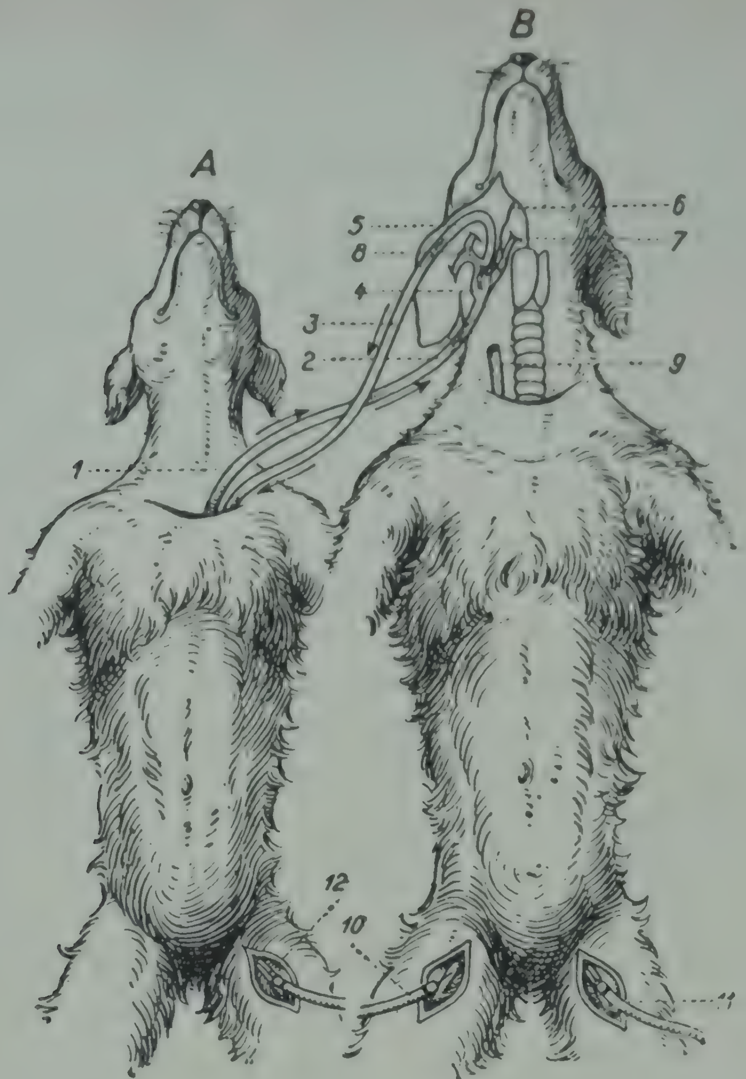


Fig. 399. Perfusion of isolated carotid sinus. A, dog supplying perfusion blood. B, dog of which the sinus is perfused. (1) Carotid artery of dog A ; (3) external jugular of dog A ; (6) nerve supply of carotid sinus of B. All other branches of the carotid sinus of B are tied off, so that the only blood entering it is that from A. (C. Heymans.)

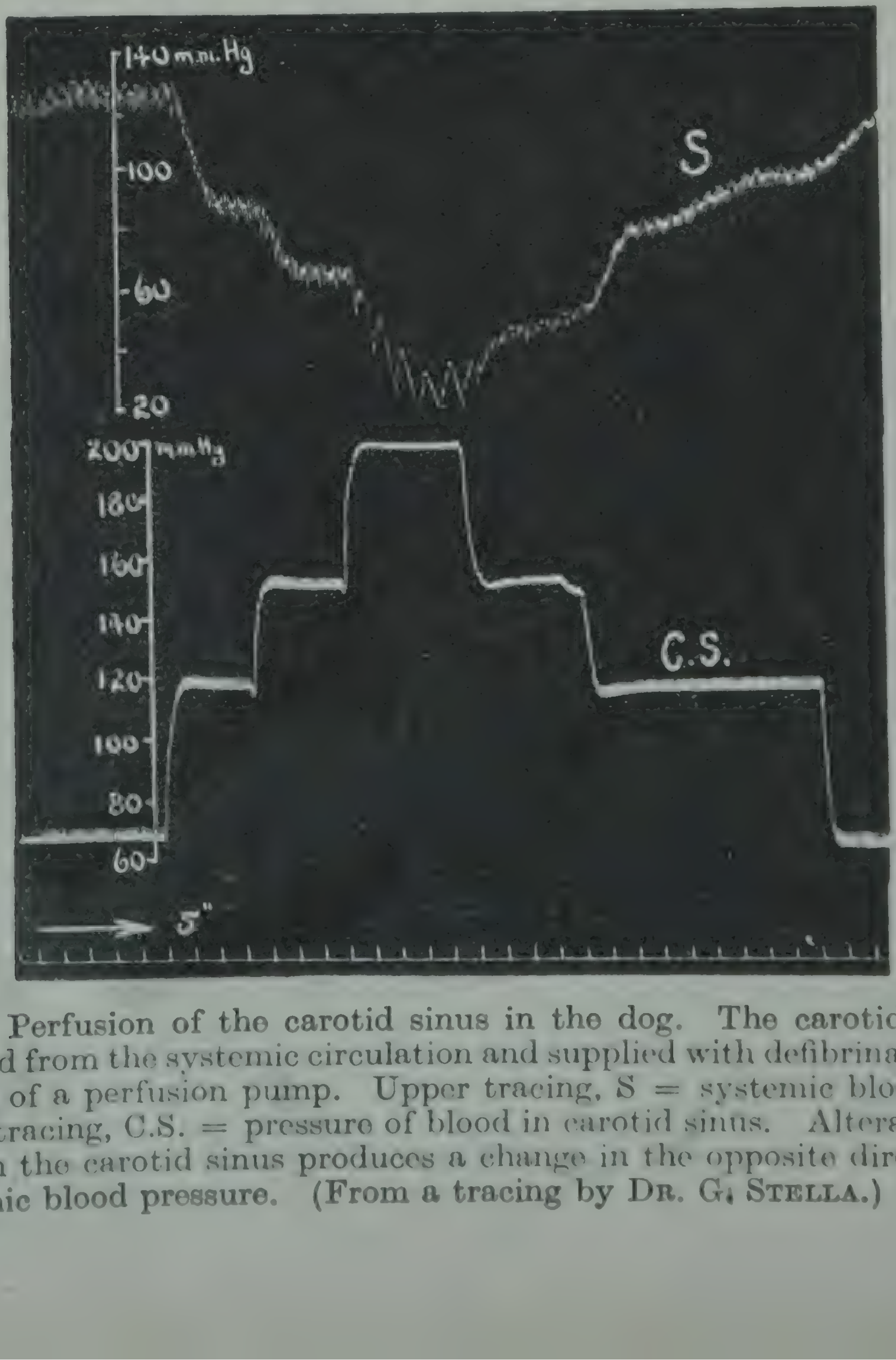


Fig. 400. Perfusion of the carotid sinus in the dog. The carotid sinuses are isolated from the systemic circulation and supplied with defibrinated blood by means of a perfusion pump. Upper tracing, S = systemic blood pressure ; lower tracing, C.S. = pressure of blood in carotid sinus. Alteration of pressure in the carotid sinus produces a change in the opposite direction in the systemic blood pressure. (From a tracing by Dr. G. Stella.)

passing along the carotid sinus segment, is returned *viâ* the external carotid artery of B to the external jugular vein of A (Fig. 399).

The pressures in the femoral artery of the recipient dog and of the perfusing supply (pump, or artery of donor) are recorded. If now the perfusing pressure is raised in any way, so that pressure is increased in the sinus of dog B, there is an immediate reduction of heart rate and fall of arterial blood pressure in the systemic circulation of the latter. If the blood pressure in the sinus of B is still further increased, there is a still further slowing of the heart and fall of pressure in dog B (Fig. 400). Conversely, by lowering the pressure in the carotid sinus of B, its systemic arterial pressure is raised, and there is at the same time cardiac acceleration. The alterations of arterial pressure in the dog B, though they accompany the alterations in heart rate, are only in part due to them, since the pressure changes still persist after elimination of the cardiac effects by section of the vagi or administration of atropine (see p. 588).

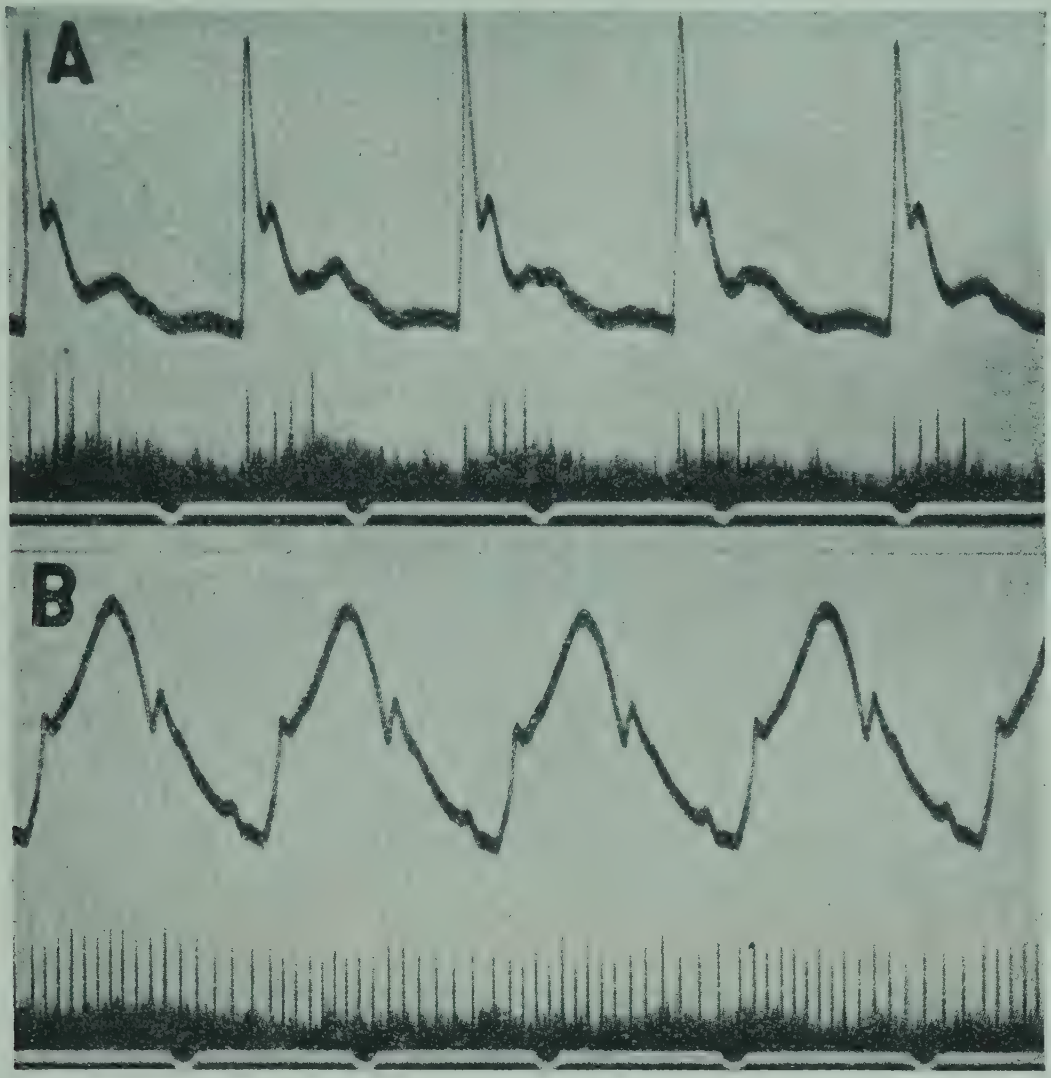


FIG. 401. Effect of increased mean blood pressure on the discharge from a single end-organ of the carotid sinus. In each tracing, the top curve is arterial pressure, middle one is electrical variations in sinus nerve, lowest one = 1/5th second. In A, B.P. = 55 mm. Hg. In B, = 135 mm. Hg. (BRONK and STELLA, *Journ. Cell. and Comp. Physiol.*, 1932, **1**, 113.)

Nerve impulses up the carotid sinus nerve are set up by rise of pressure in the carotid sinus. The electrical variations representing the nervous impulses passing along a single nerve fibre from the carotid sinus, are shown in Fig. 401. It will be seen that in tracing A, where the blood pressure was low, a burst of impulses passed along the nerve during the systolic part of each pressure-pulse curve. When the arterial pressure was raised, there was a more rapid stream of impulses, which continued also during the diastolic portion of the curve, though at a slower rate (B.).

(ii) ACCELERATOR REFLEXES. The other mechanism is a reflex, often called the Bainbridge reflex,* which is started from either auricle, and from the great veins, when these become distended by rise of venous pressure.† Afferent impulses from the auricles travel up by the vagi (perhaps chiefly by the right one) and bring about reflex cardiac acceleration by (*a*) reducing the tone of the cardio-inhibitory mechanism (the principal factor), and possibly also by (*b*) throwing the cardio-accelerator mechanism into action.

Cardiac acceleration can, of course, also result from reduction of the vagus tone. This happens when the arterial pressure falls, affecting the inhibitory zones, the carotid sinus and aortic arch.

---

* BAINBRIDGE. *J. Physiol.*, 1915, **50**, 65.

† SASSA and MIYAZAKI. *J. Physiol.*, 1920, **54**, 203.

Emotional acceleration of the heart, as, *e.g.* on hearing a loud noise, is largely due to sympathetic action, but in part to reduction of vagus tone. Then follows a fall in rate, due perhaps to intervention of the carotid sinus ; after this there may be a second rise, due to adrenaline.*

Effects of Changes in Cephalic Arterial Pressure. An important question which arises out of the discovery of the carotid sinus reflexes is that which concerns the effect on the general circulation of increased arterial pressure in the head. It has been generally agreed that an increase in the intracranial arterial pressure tends to produce, as compensatory phenomena, slowing of the heart and a fall in the somatic blood pressure.

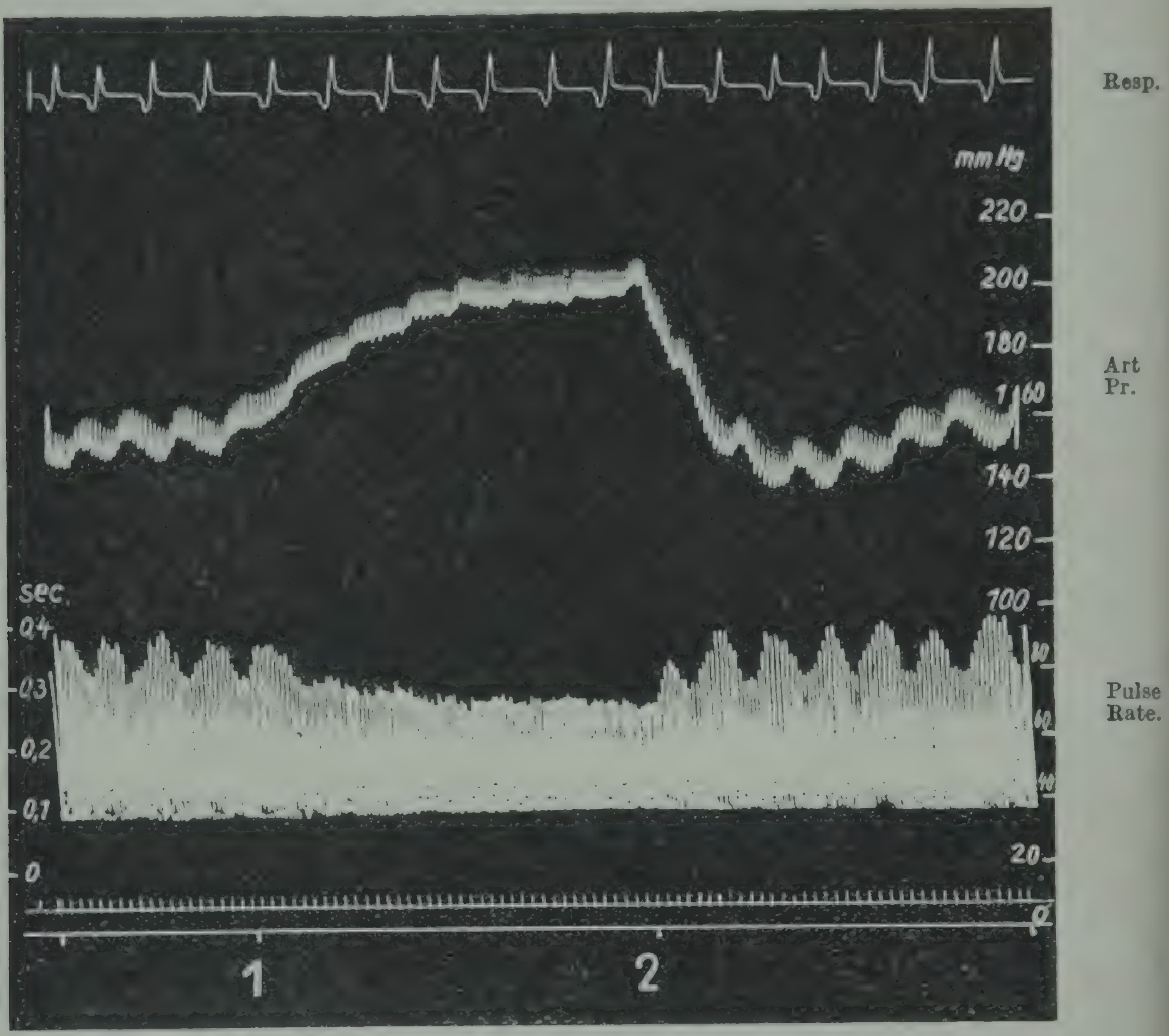


Fig. 402. Showing the effect of carotid sinus reflexes on respiration, arterial pressure, and pulse rate. In resp. trace inspiration is upwards. The right sinus and left aortic nerves have been eliminated. Between 1 and 2 the left common carotid was clipped off. Pulse rate shown by Fleisch's "time writer." Sinus arrhythmia and Traube-Hering waves are diminished during the rise of arterial pressure. Time, seconds. Ordinate on left = interval between heart beats in seconds. (Schweitzer. *Verh. dtsch. Ges. Kreislaufforsch,* 1935, **8**, 148.)

Most investigators believed that the effects were entirely due to the action of increased pressure upon the medullary centres, until it was shown by Hering that these effects disappeared when all the nerves of the carotid sinus had been destroyed. This can be illustrated by the simple experiment of lowering the cranial pressure by occluding the carotid arteries (Fig. 402). The effects disappear if the carotid sinus is denervated.

An application of the technique of the "isolated head," enables this question to be investigated. In this preparation (Fig. 403) the head of dog B is fed exclusively from dog A by anastomoses between both carotids and external jugulars of the two animals—all the tissues of the neck of B, with the exception of the two vagi, being divided. The

* Bond. *Amer. J. Physiol.*, 1943, **138**, 468.

head of B, therefore, is in vascular connection with A, and its only connection with its own body is through the two vagus nerves.

With this preparation it can be shown that a rise of pressure in A leads to a slowing of the heart in B. If the carotid sinuses of B are denervated, alterations of pressure in A no longer produce any effects upon the heart rate of B. We must conclude, therefore, that pressure changes in the head produce their effect upon the heart rate, not through any direct action on the cardio-inhibitory centre, but entirely through reflexes generated in the carotid sinus.

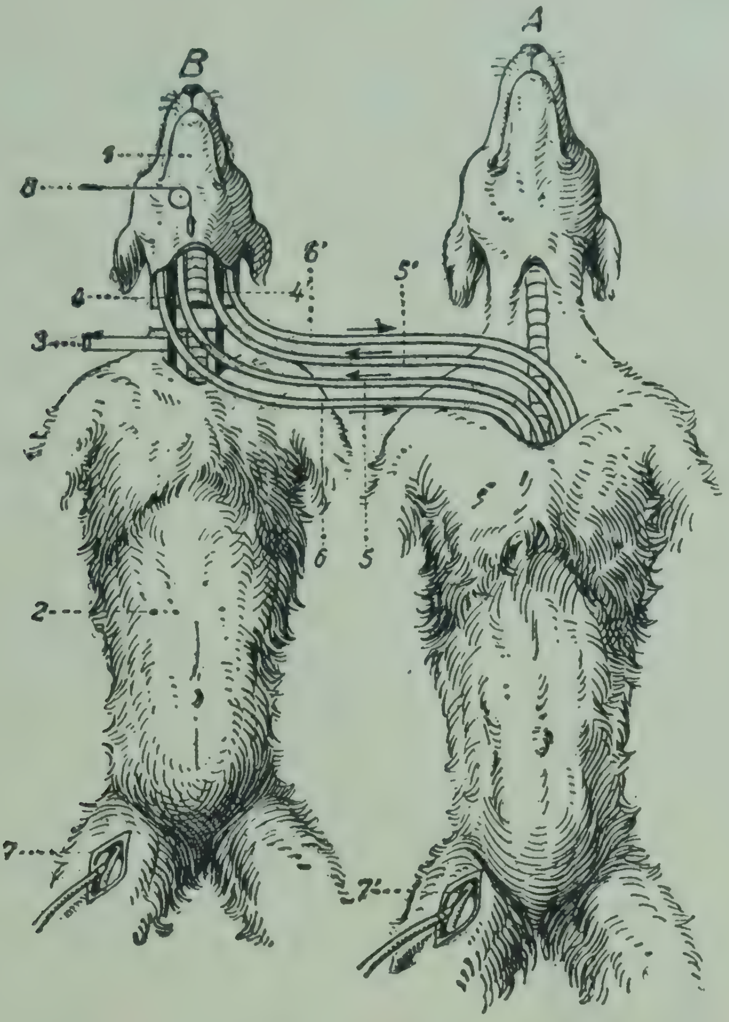


Fig. 403. Schema of the isolated head, according to J. F. Heymans. The isolated head of B is supplied with blood from dog A, and is left in connection with its own body only by the two vagus nerves. (C. Heymans.)

For Reference

Heymans, Bouckaert and Regniers (1933). "Le sinus carotidien, etc." (Doin: Paris.)

Schmidt and Comroe (1940). "Functions of the Carotid and Aortic Bodies." *Physiol. Rev.*, **20**, 115.

The Pulse Rate in Man. The normal pulse rate in adult men is about 70 per minute. It is largely influenced by bodily movements and varies considerably with age. The following table represents the average pulse rate in man at different ages :

Age in years.	Pulse rate per minute.
Fœtus	150
0	135
5	105
10	90
15	80
20	75
25–40	70
80	75

It must be remembered that marked differences in the pulse rate may be found in different individuals without having any pathological significance. Thus pulse rates of 30 per minute and 120 per minute have been observed in men who were otherwise perfectly healthy. It is slow in athletes. The pulse rate is raised by warmth and diminished by cold applied to the surface of the skin. It is also increased somewhat by the taking of food. The act of swallowing causes a reflex quickening of the rate by inhibition of the tonic vagus action.

The heart-beat is normally not perfectly regular, but shows fluctuations corresponding with the respiratory movements, the rate being accelerated in inspiration and retarded in expiration.

Irregularity of the pulse (or of any other movements) can be conveniently recorded by means of a special instrument,* which obviates the labour of measurement of the time intervals on an ordinary tracing. Fig. 402 shows a tracing made with this apparatus. The height of each stroke on the tracing shows the length of time elapsing between one pulsation and the next, so that a longer stroke means a slower pulse. The tracing shows that at inspiration there is acceleration (shorter stroke), and at expiration retardation (longer stroke). Perfect regularity of the pulse rate is abnormal. In most normal children, and in certain pathological states, the respiratory fluctuation is so pronounced as to be detectable on mere palpation; it is called *sinus arrhythmia*. Various explanations have been given for its occurrence, one being that it is due to a Bainbridge reflex from the right auricle, which is more rapidly filled during inspiration.

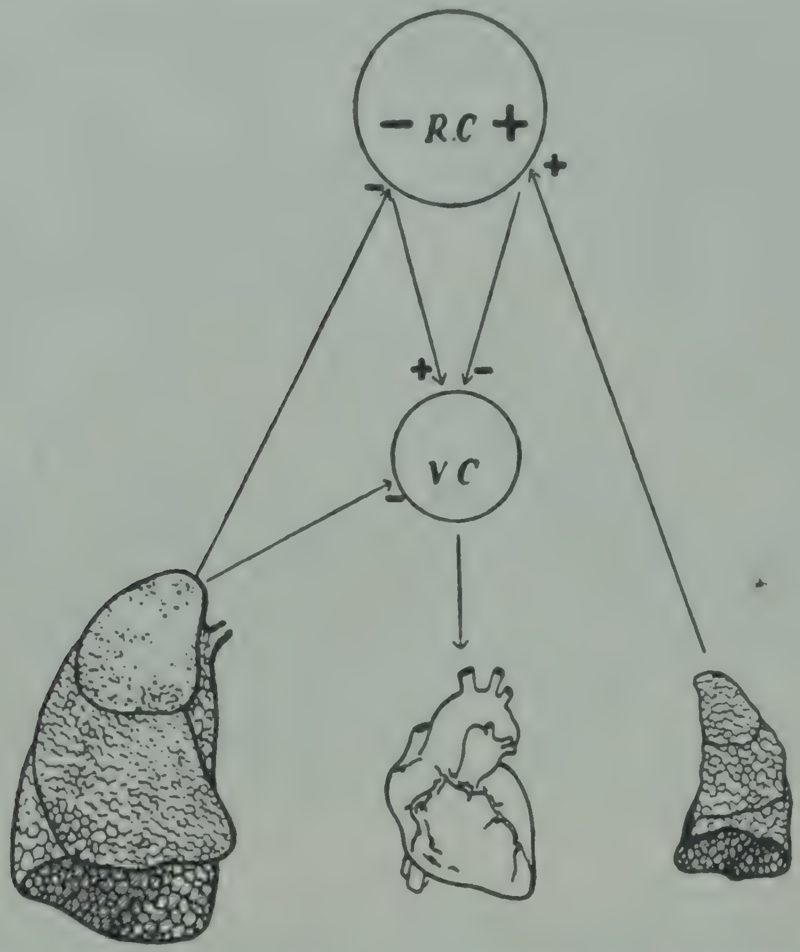


Fig. 404. Diagram showing effects of respiratory act on heart-rate. R.C., respiratory centre; V.C., vagus centre; on the right, collapse excites R.C., and indirectly inhibits V.C., so tending to produce acceleration. On left, normal inflation does the reverse, but also reflexly inhibits V.C. The effects are a resultant. (Modified from Anrep, Pascual and Rössler.)

Probably, † however, the effect is a complex one, and certainly includes some reflex effects from the lungs and some from the respiratory centre. When the lungs are inflated, the branches of the pulmonary vagus carry to the vagus centre impulses which reduce its tone; in deflation these impulses cease. Opposing these reflex effects to some extent are effects of the respiratory centre upon the vagus centre; thus inflation inhibits the respiratory centre and as a secondary result of this the vagus centre is excited. Hence the effects of the act of respiration, in addition to that due to the Bainbridge reflex, represent a resultant of excitation and inhibition at each inspiration or expiration. These effects are illustrated in Fig. 404.

NUTRITION AND METABOLISM OF THE HEART

The Coronary Circulation. The mammalian heart is nourished through the coronary arteries, the venous blood being returned to the right side of the heart. In the dog the right ventricle is mainly supplied by the right coronary artery and drained by the anterior cardiac veins, the left by the left coronary and drained by the coronary sinus.* The extent of drainage into Thebesian vessels, which open chiefly into the right ventricle, is doubtful. In the human heart the respective contributions of right and left coronary arteries varies, the most favourable (in 34 per cent.) being a balance between the two, and the least favourable (18 per cent.) one where the left artery predominates—a condition predisposing to the most serious consequences should left coronary occlusion occur,† because the larger branches of the coronary arteries show little anastomosis.‡ Hence, the occlusion of a large branch may lead to death. But anastomoses readily develop if one vessel is occluded, and are more free between veins. Capillaries are very abundant, about 11 m. of capillary per c. m.m. of heart muscle.§

---

* Fleisch, A. *Zeitschr. f. klin. Med.*, 1932, **119**, 618; Johnsen. *Acta. Physiol. Scand.*, 1945, **10**, 342.

† Anrep, Pascual and Rössler. *Proc. Roy. Soc.*, 1936, **119B**, 191, 218.

Despite its importance, we are still in doubt regarding many aspects of the coronary circulation as it operates in the normal heart in the intact chest, since nearly all available methods of experiment introduce some abnormal factor of incalculable effect.

The flow of blood through the coronary circulation may be measured by feeding an excised heart, through the coronary arteries, with blood from a pump or heart-lung preparation, and collecting the total outflow from the cut pulmonary artery. Another method is by introducing a cannula through the wall of the right auricle into the coronary sinus and collecting the blood which flows out from the latter. By a comparison of these two methods, it is found, in the dog, that the flow through the coronary sinus *usually* forms about three-fifths of the total blood passing through the coronary arteries, the remainder passing by the anterior cardiac veins and other lesser channels.|| It is therefore possible to measure the flow through the coronary sinus under varying conditions, and, by multiplying the figures so obtained by 5/3, to obtain the *approximate* total flow through the coronary circulation. In grossly abnormal conditions, *e.g.* when the heart is fibrillating, this fractionation no longer holds good.

Since the efficiency of the heart is ultimately dependent on an adequate supply of oxygen, the importance of the coronary blood flow is evident. About 5 per cent. of the cardiac output is diverted to the coronary circulation. The amount of blood passing through it per gramme of heart per minute may rise to as much as 5 c.c. In man during violent exercise the coronary flow probably amounts to about 1·5 l./min.

Factors Controlling Coronary Flow in the Denervated Heart. Other factors remaining the same, whether the work of the heart, or its rate, is varied or not, the coronary flow depends primarily on the mean pressure at which the coronary arteries are supplied. The great effect of general arterial pressure on the coronary flow is most readily demonstrated on the heart-lung preparation; this enables the mechanical factors controlling coronary blood supply to be studied alone, any influence of the nervous system being eliminated. Results of an experiment are shown in the following table :—

Heart weight, 107 grammes. Total Systemic output per minute, 1,400 c.c.

Arterial pressure mm.Hg.	Coronary circulation per minute cc.
60	50
100	90
128	124
166	208
190	500

But this circulation is also adapted to the amount of work that the heart is doing, since the oxygen usage of the heart increases with the amount of work it is given to do. When we investigate the coronary flow in the heart-lung preparation, we find, accordingly, that it is increased whenever an increase is

---

* Gregg and Shipley. *Amer. J. Physiol.*, 1947, **151,** 13.

† Schlesinger. *Blood, Heart and Circulation.* Science Press, 1940, p. 61.

‡ Schlesinger. *Amer. Heart J.*, 1938, **15,** 528.

§ Wearn. *J. Clin. Invest.*, 1924, **1,** 572.

|| Evans and Starling. *J. Physiol.*, 1913, **46,** 413.

demanded to supply the necessary oxygen. Hence we find that reduction in the oxygen saturation of the blood passing to the heart, or increase in the oxygen demand by it, soon gives rise to dilatation of the coronary vessels. This dilatation is almost proportional to the degree of anoxæmia, so that the accessibility of oxygen to the myocardium may remain unaltered, *i.e.* within very considerable limits, it is found that the consumption of oxygen by the heart is independent of the oxygen saturation of the arterial blood. The effect of increasing deoxygenation of the blood on the coronary flow is shown in the accompanying diagram (Fig. 405). The effect of altered arterial pressure therefore, even in the denervated heart-lung preparation, is not solely mechanical, but is in part due to increased oxygen abstraction, the result of increased cardiac work. This is probably to be explained as due to a direct effect of oxygen lack on the plain muscle of the coronary vessels; it is seen to occur at once when traces of cyanides are added to the blood, so as to suppress oxygen utilisation (Fig. 405). It has been ascribed to the action of metabolites produced in the heart muscle, and normally eliminated by oxygen.*

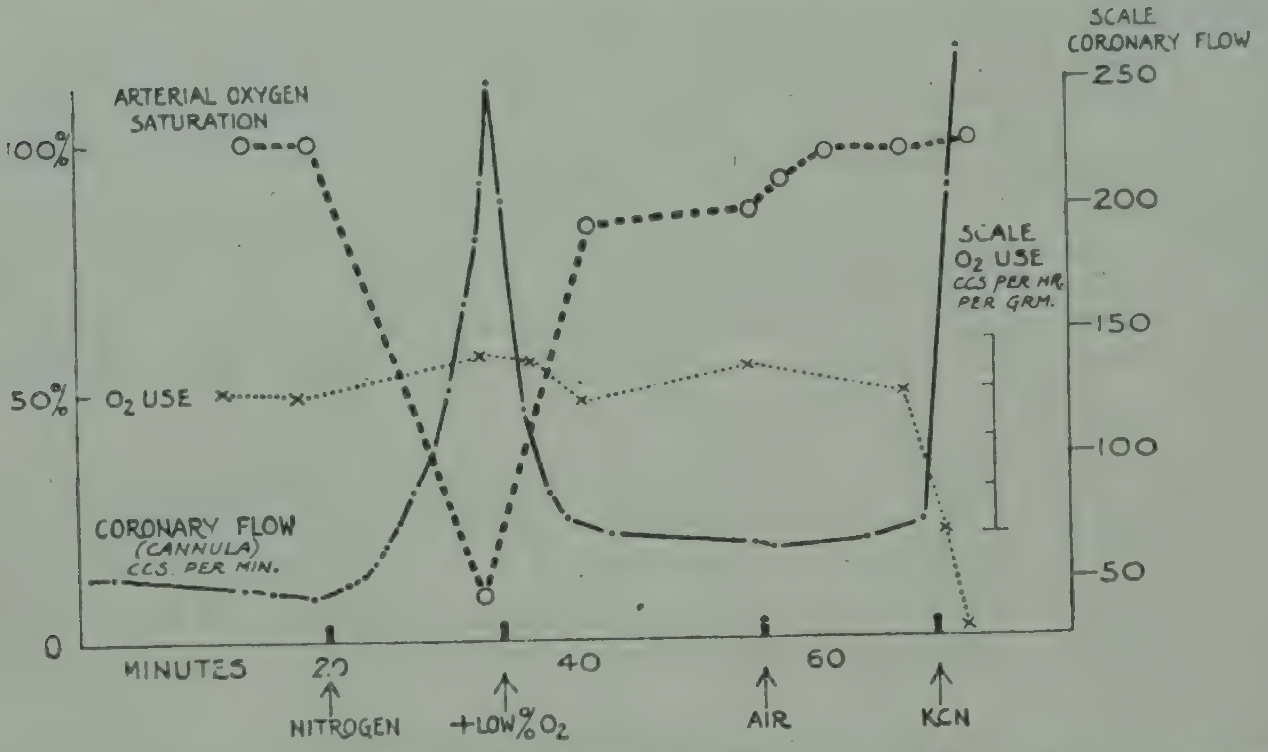


FIG. 405. Curves showing the Relation of the Coronary Blood Flow and of the Oxygen Consumption by the Heart to the Degree of Oxygen Saturation of the Blood passing through the Coronary Vessels. (HILTON and EICHHOLTZ.)

Over a moderate range, alteration in the heart-rate has little effect on the total coronary flow, but when the rate becomes very high the relative shortening of diastole results in reduction of flow.

Increase of right intraventricular pressure reduces coronary flow.

Carbon dioxide has a dilator influence on the coronary vessels, but this is slight compared with that of anoxæmia, and is probably due to the accompanying slight increase in hydrogen ion concentration.

Pitressin has powerful constrictor effects and considerably reduces coronary flow.

Control of Coronary Flow in the Innervated Heart. The coronary arteries are richly innervated from both the vagus and the sympathetic, and the vagal fibres are constrictor to the coronary vessels and normally exhibit a tonic action. Coronary vaso-constrictor and vaso-dilator reflexes probably enter into many vaso-motor reflexes, elicited from almost any part of the body.† For example, a rise of pressure in the carotid sinus probably leads to reflex coronary constriction. Section of the vagi not only quickens the heart rate but also relaxes the coronary vessels and allows an increased flow. Hence the coronary flow in the denervated heart is greater than when it is normally innervated. Further, when the output of the heart in the intact body is increased, there is, in addition to the reflex cardiac acceleration, also an independent reflex relaxation of the coronary vessels, whereby the circulation through the heart is increased : this is due chiefly to reduction of vagal tone and is not exhibited by the denervated heart-lung preparation, in which the effect of alteration of output has but little effect on the coronary flow. Experiments on unanæsthetised dogs, in which the coronary flow was measured by a thermostromuhr, have shown* that during exercise there is a large increase in coronary flow : sympathectomy did not affect this response, but double vagotomy, though it resulted in an increase of the resting flow, abolished the increase on exercise.

---

* KATZ and LINDNER. *Amer. J. Physiol.*, 1939, **126**, 283.

† GREENE. *Amer. J. Physiol.*, 1935, **113**, 399.

On stimulating the sympathetic nerve supply to the heart, or injecting adrenaline, there is an increased frequency and energy of contraction and the increased beat is always associated with greatly increased coronary blood flow. It has been inferred that the sympathetic nerves to the heart have a direct dilator effect upon the coronary arterioles,† but owing to the concurrent effect on the myocardium, with consequent increased production of metabolites, it is still uncertain whether this is so.

The above factors may be all involved in the adaptation of the blood flow to the needs of the heart when the activity of this organ is increased, either by increased arterial resistance or by increased output, in both of which the oxygen usage of the heart is increased.

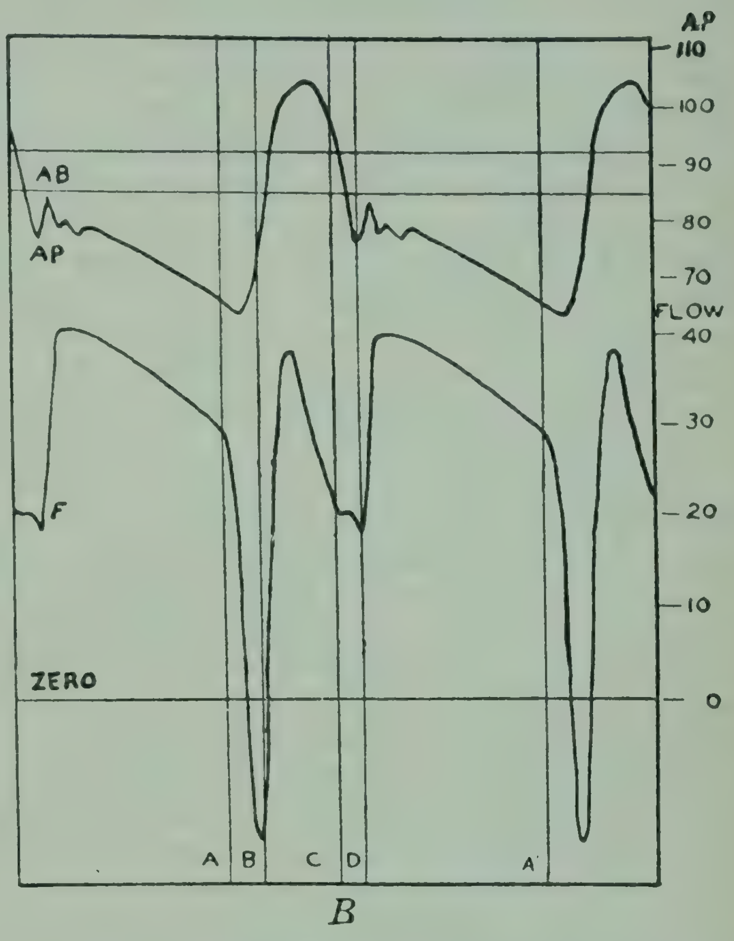


FIG. 406. Rate of inflow F into a branch of a left coronary artery, and aortic blood pressure AP, in dog. A = onset of systole. B = ejection phase begins. C = end of systole. D = closure of aortic valves. (GREGG and GREEN. *Amer. J. Physiol.*, 1940, **130**, 114.)

Phasic Changes during Cardiac Cycle. The influence of the various phases of the cardiac cycle upon the coronary flow has been the subject of much debate. *A priori* we should expect that in systole blood would be squeezed out of the cardiac veins, and prevented from entering by the arteries. Anrep studied the coronary inflow and outflow separately, and found that, in fact, these do not run parallel. The coronary outflow shows three waves, very similar to those seen in the intra-auricular pressure ; the most rapid flow is seen during ventricular systole, and is no doubt due to a massaging action of the ventricular contraction expelling blood from the coronary veins. The other waves of the outflow are probably due to the effects of auricular contraction.

Study of the rate of inflow into the coronary arteries by the most critical methods shows that the most rapid entry is during diastole. During the stage of isometric ventricular contraction the inflow is suddenly reduced, or even reversed momentarily. As soon as the ejection phase begins, the coronary flow increases again and roughly follows the rise and fall of aortic pressure until the end of systole. After the dicrotic notch, with ventricular relaxation a sudden increase of inflow occurs, and this then slowly falls off with declining aortic pressure, until onset of the next systole. (Fig. 406).

* ESSEX *et al.* *Amer. J. Physiol.*, 1943, **138**, 687.

† GREENE. *Amer. J. Physiol.*, 1935, **113**, 361.

For Reference

ANREP (1926). The Regulation of the Coronary Circulation. *Physiol. Rev.*, **6,** 596.
GREGG (1946). The Coronary Circulation. *Physiol. Rev.*, **26,** 28.
WEARN (1939). Morphological and Functional Alterations of the Coronary Circulation. (*Harvey Lects.*, 1939–40, p. 243).

Metabolism of Cardiac Muscle. The energy for the contraction of cardiac muscle comes ultimately from oxidative processes, and we have seen that the oxygen utilisation of the heart depends on the work done, and on its diastolic volume.

When the work of the *isolated* (*i.e.* denervated) heart is increased by raising either the aortic pressure or the output, the oxygen usage rises, and usually the efficiency rises too. The two methods of raising the work in the heart-lung preparation do not, however, produce the same effects on oxygen usage and hence on efficiency; when work is raised by increase of resistance the oxygen usage shows a greater increase than it does when the work is raised to an equal extent by increase of output. Hence a greater mechanical efficiency results from a given increase of work by increase of output than by increase of resistance. These inferences need to be qualified, however, for the heart *in situ.*

When the work of the heart *in situ* in the normal body is raised, the heart rate is usually altered: *e.g.* when the work is increased by raising the arterial resistance the heart rate is slowed (Marey's reflex), while if it is raised by increase of output the rate is accelerated (Bainbridge reflex). These results are due to alterations, in the appropriate direction, of vagus and sympathetic tone. The question as to what happens to oxygen usage and efficiency when the work of the heart *in situ* is raised therefore resolves itself into two subsidiary questions, viz. (1) What is the effect of alteration of rate ? (2) What, if any, is the effect of a change of vagus and sympathetic tone, apart from that due to change of heart rate ?

It is generally admitted that increase of rate leads to an increase, though not a proportionally great one, in metabolism. When the heart is slowed by increase of vagal tone there is a greater reduction in the oxygen usage than could be accounted for by the reduction in number of beats, and conversely increase of the sympathetic tone (or the action of adrenaline) causes a disproportionately great increase in metabolism. Hence it was found that, owing to the increased vagal tone, increase of work due to rise of arterial pressure in the innervated heart was often effected without any increase of oxygen usage at all, and so the heart worked *much more efficiently at the higher load than at the lower.* This provided a marked contrast to the denervated heart, which was very well seen in some cases where for some reason slowing did not occur. The adjustment is called the "unloading reflex."

A further contrast is seen when comparison is made between increase of work by increased resistance and increased output respectively, *e.g.* in the following example (see table below) the work was doubled.

These results teach us two things. First, that it is a property of cardiac muscle *per se* to respond by a greater increase of metabolism to rise of arterial pressure than to increase

Denervated Heart. (Evans and Matsuoka, 1914.)

Method of Increasing Work.	Ao. Pressure min. Hg.	Output c.c./min.	Rate per min.	Work kg.-m./min.	O_2 used c.c./min.	Efficiency. Per cent.
Pressure raised .	40	282	130	0·18	1·63	5·5
	80	280	130	0·34	1·98	8·4
	160	280	130	0·68	4·16	7·9
Output raised . .	80	282	130	0·35	2·33	7·3
	80	562	130	0·78	2·64	14·2
	80	1,350	130	2·91	4·96	28·2
Innervated Heart. (Gollwitzer-Meier *et al.*, 1938.)						
Pressure raised .	71	550	125	0·62	3·5	8·8
	136	576	97	1·25	3·2	19·5
Output raised . .	60	630	154	0·60	3·8	7·9
	63	1,128	156	1·28	5·4	11·8

of output and to reach the limits of its performance earlier in the former case than in the latter. Secondly, that when provided with its normal innervation, so that cardiac reflexes are operative, this property is so modified that rises of arterial pressure can be dealt with at least as economically as increases of output.

The cardiac metabolism also depends on the temperature, and on the influence of certain chemical agencies. Thus, if adrenaline is present there is a large increase in the oxygen usage of the heart, which is made possible by the accompanying augmentation in the coronary blood flow. The following experiment illustrates this effect :—

EFFECT OF ADRENALINE 0·2 MG. ON HEART-LUNG PREPARATION OF DOG. (EVANS and OGAWA. *J. Physiol.*, 1914, **47,** 446.)

Heart 58·5 g ; art. press. = 100 mm. Hg ; output c̄. 10·5 l./hr.

	Period of Observation mins.	Average Heart Rate p. m.	cc. p. hour. by H.-L. P.		R.Q.
			O_2.	CO_2.	
Before adrenaline .	80	145	202	164	0·81
After ,, .	30	218	597	463	0·78

A small dose of adrenaline may enable either a normal or a hypodynamic heart to perform tasks previously too great for it. The heart volume is reduced and much greater pressures and volumes of blood can be dealt with. But as the oxygen usage shows, the effect, at least on the *denervated heart,* is not economical but profligate. The heart has been flogged to inordinate activity, and although the work has been done the cost has been high and the efficiency correspondingly low. In fact, by prolonging the administration, the heart works apparently well for a time, and then, with scarcely any warning, passes into acute failure.

When adrenaline is given to an *innervated heart* connected by its nerves to a separately perfused head, the " unloading reflex " already mentioned is brought into play, the increased vagus tone opposing the stimulating effect of adrenaline on the cardiac metabolism, which shows a sharp peak followed by a fall ; the coronary flow is increased greatly so that the venous blood is never very deoxygenated.

Adrenaline also exerts an effect owing to its action on centres in the brain. To study the central effect of adrenaline, the drug is given to the perfused head of an innervated heart-lung preparation, but not to the heart itself. The result is great reduction of metabolism, slowing and transitory coronary constriction.

Finally, when adrenaline is given to the whole animal we have both types of action—the unloading reflex from the heart itself and the central unloading reflex, both of which antagonise the direct effect of adrenaline on the heart. As the outcome of the competition between them there usually results an actual *diminution* of oxygen usage and of rate, while the work done may even increase. Coronary flow is greatly augmented, and the coefficient of oxygen utilisation is much reduced, because the direct effect of adrenaline on the coronary vessels is greater than the reflex constrictor effect. The efficiency is greatly increased, as contrasted with the lowered efficiency in the isolated heart.

In chemical composition the heart muscle somewhat resembles striated muscle ; it contains glycogen, adenosine triphosphate, and creatine phosphate, though the latter is in smaller amount than in striated muscle. It would be expected, therefore, that it might resemble striated muscle in the nature of the chemical reactions associated with contraction, but, in fact, there are certain clear differences.

The metabolic changes in the heart can conveniently be studied by analysis at intervals of the blood in the heart-lung preparation, and this has often been done. But it has the drawback that the metabolism of the lungs cause changes in the blood which are by no means negligible ; the most conspicuous change introduced by the lungs is a great acceleration of glycolysis, by which the blood glucose is converted into lactic acid in passing through the lungs.*

---

* EVANS, HSU and KOSAKA. *J. Physiol.*, 1934, **82,** 41

In order to eliminate the lungs, and so obtain a clearer picture of the sugar usage and metabolism of the blood-fed heart doing work, the lungs can be replaced by an artificial oxygenator. Such a heart-oxygenator preparation, * which has other applications also, is shown in Fig. 407.

The two auricles are supplied with blood through the adjustable taps R and L, and cannulæ 1 and 2, from a narrow common reservoir V.R. This reservoir is kept full to overflowing with blood from a stock reservoir, S.R. by the pump P, the blood passing on its way through a warming tube W and a filter F. The unwanted overflow from V.R. passes by the siphon S along the overflow tube O.T., back again to the stock reservoir. Blood expelled from the aorta passes by cannula 4 through an adjustable pneumatic resistance A.R. (as used for a heart-lung) back into V.R.; the rate of output can be measured when required in the graduated vessel A.O. (Not all the blood expelled by the left ventricle passes out by cannula 4, because the coronary supply has been deducted from it.) Blood from the right ventricle (which consists of that entering from cannula 1 plus that returning from cardiac veins) passes through the variable resistance P.R. (which can be omitted), through the graduated vessel P.O., and on to the spinning plate of a Drinker oxygenator, D.O. This is filled with suitable gas, generally oxygen containing 5 per cent. CO_2, and the blood flowing down its sides becomes arterialised and enters the stock reservoir S.R. again. Auricular and arterial pressures are given by the four manometers shown in the figure.

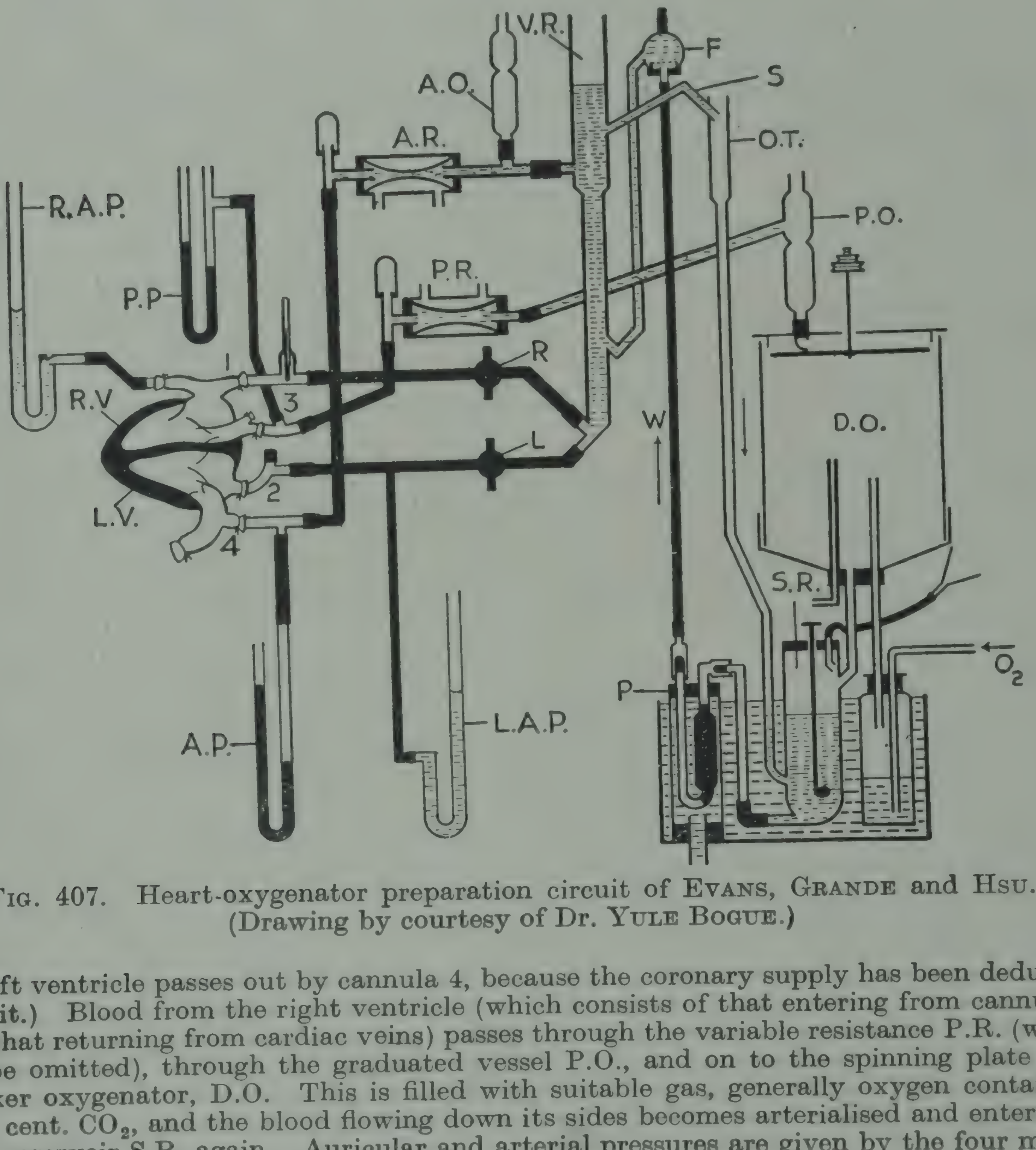


FIG. 407. Heart-oxygenator preparation circuit of EVANS, GRANDE and HSU. (Drawing by courtesy of Dr. YULE BOGUE.)

With this preparation the isolated heart can perform work at reasonable rates for several hours, the conditions of pressure and output on the two sides being independently adjustable. It will also be evident that by closing tap R, the only blood entering and leaving the right side will be the coronary venous blood, the rate of flow or chemical composition of which can thus easily be determined.

It can be calculated that if the heart were only oxidising its own glycogen as a source of energy and not replacing it, a dog's heart containing about 0·7 per cent. glycogen should only be able to work for at most two hours, before exhausting all its glycogen. Since the isolated heart easily works for four or five hours, at the end of which time it often contains an amount of glycogen normal for a fresh heart, it is evident that it either does not normally use glycogen, or else that the glycogen is replaced as used. Now it has long been

---

* EVANS, GRANDE and HSU. *Q. J. Exp. Physiol.*, 1934, **24**, 283.

known that the heart can remove sugar from the blood, but it has also been shown that the amount used under ordinary conditions is too small to supply the energy requirements.* It does, however, quite unlike skeletal muscle, readily use lactic acid from the blood.*† It also easily consumes pyruvic acid.‡ The amounts of glucose and lactate used vary reciprocally, so that if lactate is absent, glucose is used in greater amount, and *vice versâ*. The ordinary dog heart, in a heart-oxygenator circuit uses about 70 mg. glucose and 200 mg. lactic acid per 100 g./hour,† but if the work is increased the amounts of both are greater. §

The lactic acid used by the heart is normally formed from blood glucose by glycolysis in the blood stream, this process being accelerated in passage of the blood through the lungs. If, in a heart-oxygenator circuit, the heart is allowed to use up all the glucose and lactate, or if it is subjected to intense work, it then draws upon its own glycogen. It should be noted that in the absence of lactate from the blood, the heart becomes weak, but is restored to a better condition by the addition of lactate, which is more effective for the purpose than is glucose. The glycogen of heart muscle is formed from blood glucose, but not from lactate.‖

In addition to carbohydrate and lactic acid, the heart muscle probably utilises fat¶ ; its respiratory quotient is about 0·8 to 0·85. It is uncertain at present whether the heart draws on its glycogen normally, and continually replaces it from glucose, or whether it normally burns glucose and lactate direct, keeping the glycogen in reserve for emergencies. In oxygen lack, lactic acid is formed by the heart, and glycogen is lost.

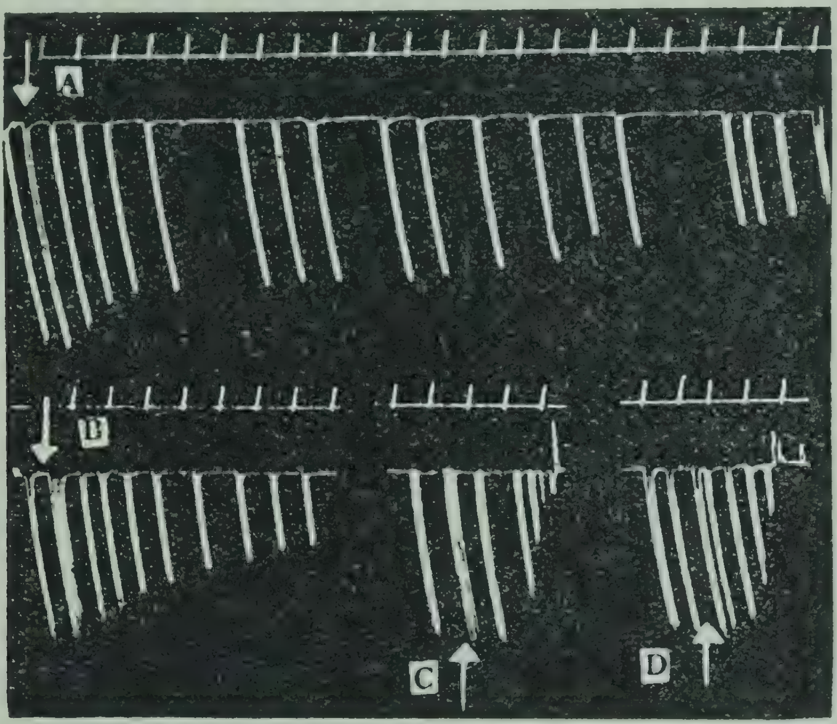


FIG. 408. Illustrating the effect of asphyxia on the frog's heart. Isometric contractions were taken for short periods at irregular intervals—downstroke = systole, time = minutes. The arrow shows commencement of full anaerobiosis.
A. Normal ventricle filled 0·3 c.c. of equal parts plasma and Ringer's fluid.
B. Normal ventricle filled 0·3 c.c. unbuffered Ringer's fluid.
C. Ventricle after poisoning with iodoacetic acid, 1 part in 20,000. Filling as in B.
D. Ventricle poisoned iodoacetic acid filled with 2 c.c. equal parts plasma and Ringer's fluid. (CLARK, *Quart. J. Exp. Physiol.*, 1935, **25**, 167.)

Further light is thrown on the chemistry of cardiac muscle by the study of the effect of oxygen lack and of iodoacetic acid. It will be recalled that when a skeletal muscle is poisoned with iodoacetic acid, no lactic acid is formed on contraction, but that contractions can be repeated in the presence or absence of oxygen until all the creatine phosphate has been broken down.

* McGINTY and MILLER. *Amer. J. Physiol.*, 1933, **103**, 712.
† EVANS, GRANDE and HSU. *Q. J. Exp. Physiol.*, 1935, **24**, 348.
‡ BRAUN-MENENDEZ, CHUTE and GREGORY. *Quart. J. Exp. Physiol.*, 1939, **29**, 91.
§ BOGUE, EVANS, GRANDE and HSU. *Q. J. Exp. Physiol.*, 1935, **26**, 213.
‖ BOGUE, EVANS and GREGORY. *Quart. J. Exp. Physiol.*, 1937, **27**, 27.
¶ CRUICKSHANK and KOSTERLITZ. *J. Physiol.*, 1941, **99**, 208.

The heart, poisoned with iodoacetic acid, can contract for an indefinitely long time if oxygen is available * ; its usage of lactate is unaltered and its phosphagen content not reduced. When oxygen is withdrawn from it, however, it fails after a very few beats † (Fig. 408). We may provisionally interpret these results as follows : The energy for contraction is given immediately or mediately by the breakdown of A.T.P. and phosphagen ; the energy for resynthesising these is provided normally by the *oxidation* of pyruvic or lactic acid. If oxygen is lacking, however, glycogen is broken down instead (as is normal in all cases for skeletal muscle), with formation of lactic acid, and this change now provides the necessary energy. When poisoned with iodoacetic acid, formation of lactic acid from glycogen is no longer possible ; hence, withdrawal of oxygen leads to rapid failure of the heart.

It seems that this oxidation of pyruvic and lactic acid represents an essential aspect of the metabolism of the heart. If the lactate is present in abundance in the blood it is that which is burned. When adrenaline is given the lactate demand of the heart is greatly increased. If lactate is lacking from the blood, then it must be obtained by breakdown of the heart's own glycogen. This, however, is either at once, or later on, replaced from blood sugar if any is available ; but if not, the glycogen will continue to fall until, when it is all, or nearly all, exhausted, the heart fails.

This theory, that oxidation of lactic acid is the central point of cardiac metabolism, explains why lactate cannot form glycogen, but why when abundant in the blood it can prevent glycogen loss.

It may well be that the fundamental difference between cardiac and skeletal muscle is to be sought in two structural differences. First the blood supply of cardiac muscle is relatively greater, and secondly the cardiac muscle cells have no sarcolemma. The first provides for a much more abundant oxygen supply, thereby facilitating all oxidative arrangements ; the second enables lactate to pass readily into or out of the muscle cell and so, if present in the blood, to be readily available for the latter. The ability of the heart to utilize lactate places it in a favourable position in severe muscular exercise.

For Reference

CRUICKSHANK (1936). Cardiac Metabolism. *Physiol. Rev.*, **16,** 597.
CLARK *et al.* (1938). " The Metabolism of the Frog's Heart." (Edinburgh and London : Oliver and Boyd.)
EVANS (1939). Metabolism of the Heart. *Edinb. Med. J.*, N.S. **46,** 733.

Books and Papers for Reference (General)

HEYMANS (1929). Le Sinus Carotidien. (Louvain, Paris and London : H. K. Lewis.)
McDOWALL (1938). " The Control of the Circulation of the Blood." (London : Longmans.)
WIGGERS (1923). Modern Aspects of the Circulation in Health and Disease. (Philadelphia and New York : Lea and Febiger.)
WIGGERS (1928). The Pressure Pulses in the Cardiovascular System. (London : Longmans, Green.)

---

* CLARK, EGGLETON and EGGLETON. *J. Physiol.*, 1931, **72,** 25P.

† CLARK. *Q. J. Exp. Physiol.*, 1935. **25,** 213.

CHAPTER XXXIII

DYNAMICS OF THE CIRCULATION

BLOOD PRESSURE

Since the blood travels from the arteries to the veins, it is evident that the pressure which drives it onwards must get less and less as it passes along.

On cutting through an artery, blood escapes from the end nearest the heart, with great force and in a series of spurts, each of which corresponds to a contraction of the ventricles. This shows that in the arteries the blood is at a high pressure, and that the flow from the heart is a pulsatory one.

The same lesson was learnt by an experiment, which was first performed in 1732 by the Rev. Stephen Hales,* and may be described in his own words :

> " In December I caused a mare to be tied down alive on her back. Having laid open the left crural Artery about three inches from her belly, I inserted into it a brass Pipe, whose bore was one sixth of an inch in diameter ; and to that, by means of another brass Pipe which was fitly adapted to it, I fixed a glass Tube, of nearly the same diameter, which was nine feet in length : Then untying the Ligature on the Artery, the blood rose in the Tube eight feet three inches perpendicular above the level of the left Ventricle of the heart : But it did not attain to its full height at once ; it rushed up about half way in an instant, and afterwards gradually at each Pulse twelve, eight, six, four, two, and sometimes one inch : When it was at its full height, it would rise and fall at and after each Pulse two, three, or four inches."

Arterial and venous blood pressures are now measured by special manometers, arterial pressures being recorded in terms of the pressure of a column of mercury.

If a vein be ligatured, it swells up on the distal side of the ligature. If the vein be cut across, blood escapes chiefly from the peripheral end, and instead of spurting out with each heart beat, it flows steadily and with very little force, so that light pressure by a bandage is sufficient to restrain the hæmorrhage. Since venous pressure amounts to only a few millimetres Hg., it is customary to use manometers filled with an aqueous solution, *e.g.*, sodium citrate, for venous pressure experiments.

By taking the pressure at different parts of the circulation, we obtain a distribution which is represented roughly in Fig. 339, p. 538. The blood pressure, which is about 100 to 120 mm. Hg in the large arteries near the heart, falls only slowly in these arteries, so that in the radial artery it is not very much below that in the aorta. Between the medium-sized arteries and capillaries there is a very extensive fall of pressure as the blood passes through the arterioles, so that in the capillaries the pressure on an average may be taken as 10 to 20 mm. Hg ; from the capillaries to the veins the blood pressure falls steadily until in the big veins near the heart it may be " negative," *i.e.* below atmospheric pressure.

The following table may serve to give an idea of the average height of the *mean* (not systolic) blood pressure at different parts of the vascular system in a young adult man, in the horizontal position. The pressures are all subject to considerable variations according to the activity of the individual and of the various parts and organs of the body :

Large arteries (*e.g.* carotid)		90 mm. mercury (65–110).
Medium arteries (*e.g.* radial)		85 mm. ,,
Capillaries	about	10 to 30 mm. mercury.
Small veins of arm		9 mm. mercury.
Portal vein (estimated)		10 mm. ,,
Inferior vena cava (estimated)		3 mm. ,,
Large veins of neck	from	0 to − 8 mm. mercury.

---

* Hales. " Statical Essays ; Hæmastatics, etc.," 1769, (London), **2**, p. 1.

The Arterial Pressure. The *mercurial manometer*, used for measuring and recording arterial pressure, consists of a U-tube with two vertical limbs about a foot in height, which is half-filled with mercury (Fig. 409). On the surface of the mercury of one limb is a float of vulcanite from which a stiff fine rod rises, bearing on its upper end the writing-point which writes on the travelling paper surface. Since the mercury is displaced equally, up in one limb and down in the other, it is obvious that any displacement as recorded must be multiplied by two to obtain the actual pressure in terms of a mercury column. The other limb of the manometer is connected by a side-branch and a flexible inextensible tube with a cannula, which is tied into the central end of an artery, a clip being previously placed on the artery so as to prevent the escape of blood. This limb is also connected, through a tap, with a reservoir bottle placed several feet above it, and filled with sodium citrate (4 per cent.) First, the tap is opened and the manometer, connecting-tube and cannula filled with this solution. The tap is then closed so that the manometer remains in connection only with the cannula, the mercury in that limb being now, say, about 150 mm. above that in the other. The clip is then taken off the artery. The pressure in the cannula being greater than that in the artery, a small amount of the fluid used to fill the tubes runs into the circulation. The mercury in the manometer drops to a height, usually of 100 to 120 mm. Hg., and stays about that level, rising and falling slightly with each heart beat (Fig. 410). Any blood which enters the cannula does not clot for a considerable time, owing to its admixture with the citrate.

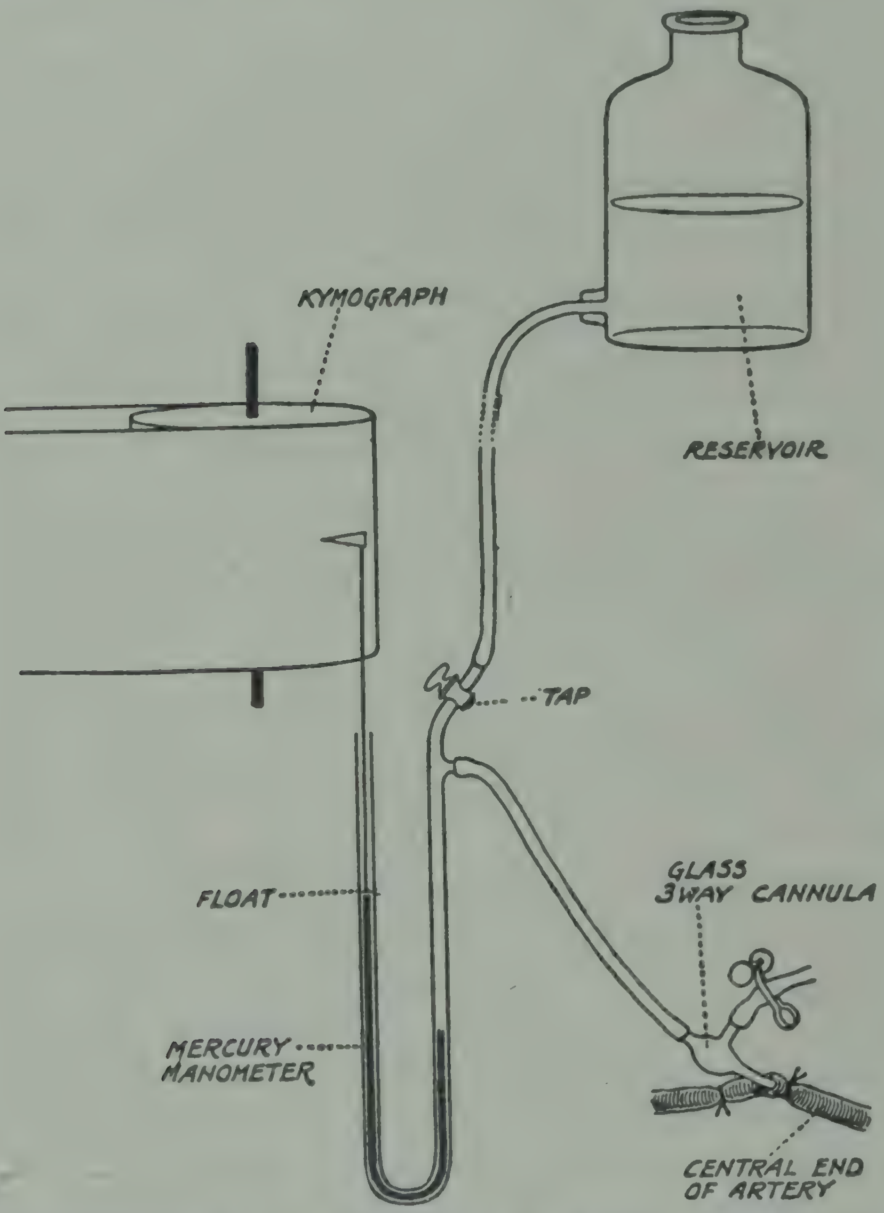


FIG. 409. Arrangement of an Apparatus for taking Blood Pressure Tracing.

The arterial blood pressure as recorded by a mercurial manometer exhibits a series of pulsations, each corresponding to a heart beat (Fig. 410). These pulsations are due to the fact that at each systole, blood is pumped out through the aortic valve, so that the artery becomes distended. During diastole, the aortic valves are closed, and blood passes on from the distended arteries into the capillaries and veins, so that the blood pressure falls. The mercurial manometer, however, owing to the inertia of the column of mercury, does not register accurately these rapid changes of pressure in the artery. With a wide-bore manometer, in fact, the oscillations may be almost imperceptible. But it gives a true and valuable record of the *mean arterial pressure*.

In order to determine the true course of the pressure changes in the arteries it is necessary to diminish to the utmost extent the inertia of the moving parts of the recording instrument, as in one of the forms of manometer already described, (p. 557) in which the pressure changes are registered photographically.

Examination of records obtained by these superior methods (Fig. 410) shows that the variation in the arterial pressure at each heart beat is much greater than would be anticipated from the tracing given by the mercurial manometer. The highest pressure, which occurs while blood is passing from the heart into the aorta, is called the *systolic arterial pressure :* the pressure at the end of diastole is the *diastolic pressure :* and the range between these two extremes is known as the *pulse pressure.* In many animals the systolic pressure as measured in the carotid artery, is under normal conditions about 120 mm., while the diastolic pressure is only about 75 mm., so that the pulse pressure is about 45 mm. Hg. As we pass outwards towards the periphery the pulse pressure becomes less and less, until finally, in the capillaries and veins, there are usually no pulse waves perceptible.

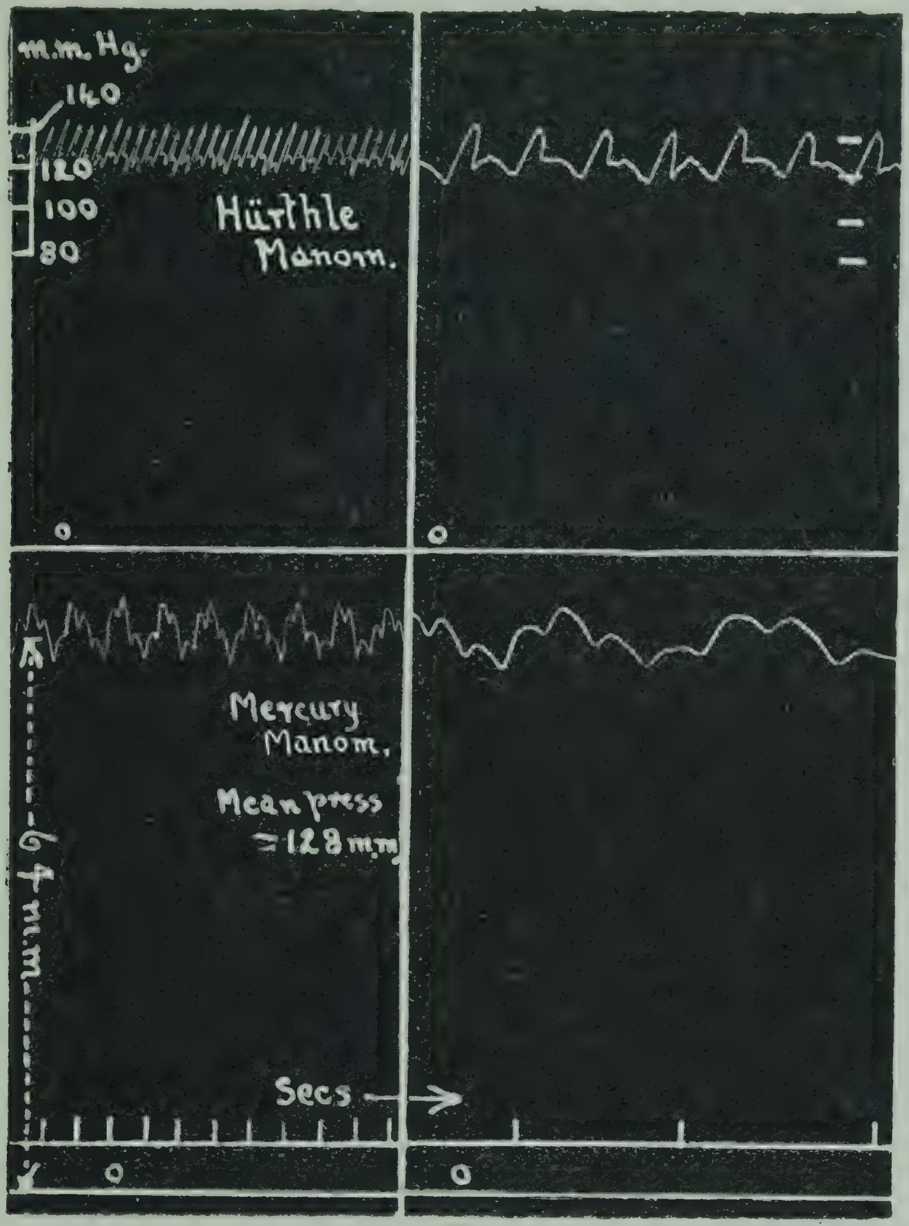


FIG. 410. Arterial Blood Pressure from carotids of Dog, recorded simultaneously with Hürthle Manometer (upper tracing) and Mercury Manometer (lower tracing), on slowly and rapidly moving surfaces. The form and extent of each pulse is much more correctly shown by the Hürthle Manometer, but the slower changes due to respiration are more evident with the Mercury Manometer. 0 = line of zero pressure. The left-hand half is on a slowly moving, and right hand on rapidly moving surface. (C. L. E.)

The Venous Pressure. In experimental animals the venous pressure may be recorded by connecting a branch of the required vein with a manometer filled with $NaHCO_3$ solution or 4 per cent. citrate, and provided with a light recording float. It is important that there should be no valve in or near the branch selected ; consequently the choice of veins is restricted, and the significance of results often in doubt. Venous pressures near the periphery are not definitely pulsatile, but in the great veins there are considerable pulsations, the venous pulse (p. 640), and these need special methods of recording.

ARTERIAL AND VENOUS BLOOD PRESSURE IN MAN

Arterial Pressure. The determination of the systolic pressure in the arteries is easily carried out by the use of the sphygmomanometer,* of which there are many modifications. This apparatus (Fig. 411) consists of a leather or canvas cuff about 10 cm. wide, which can be buckled closely round the upper arm. Inside this band is a rubber bag of the same shape, which communicates by a rubber tube with a mercurial manometer or other pressure gauge, and by a side tube with an air pump, and with an outlet valve. The band is buckled round the upper arm and the fingers of the observer are placed on the radial pulse. The bag is then distended with air so that it exercises a pressure on the arm, the pressure being indicated on the mercurial manometer. Air is forced in until the radial pulse disappears. By means of the outlet valve the air is then let slowly out of the bag until the radial pulse is just perceptible. The height of the mercurial manometer at this moment is equal to the systolic pressure in the main arterial trunk from which the brachial artery takes origin.

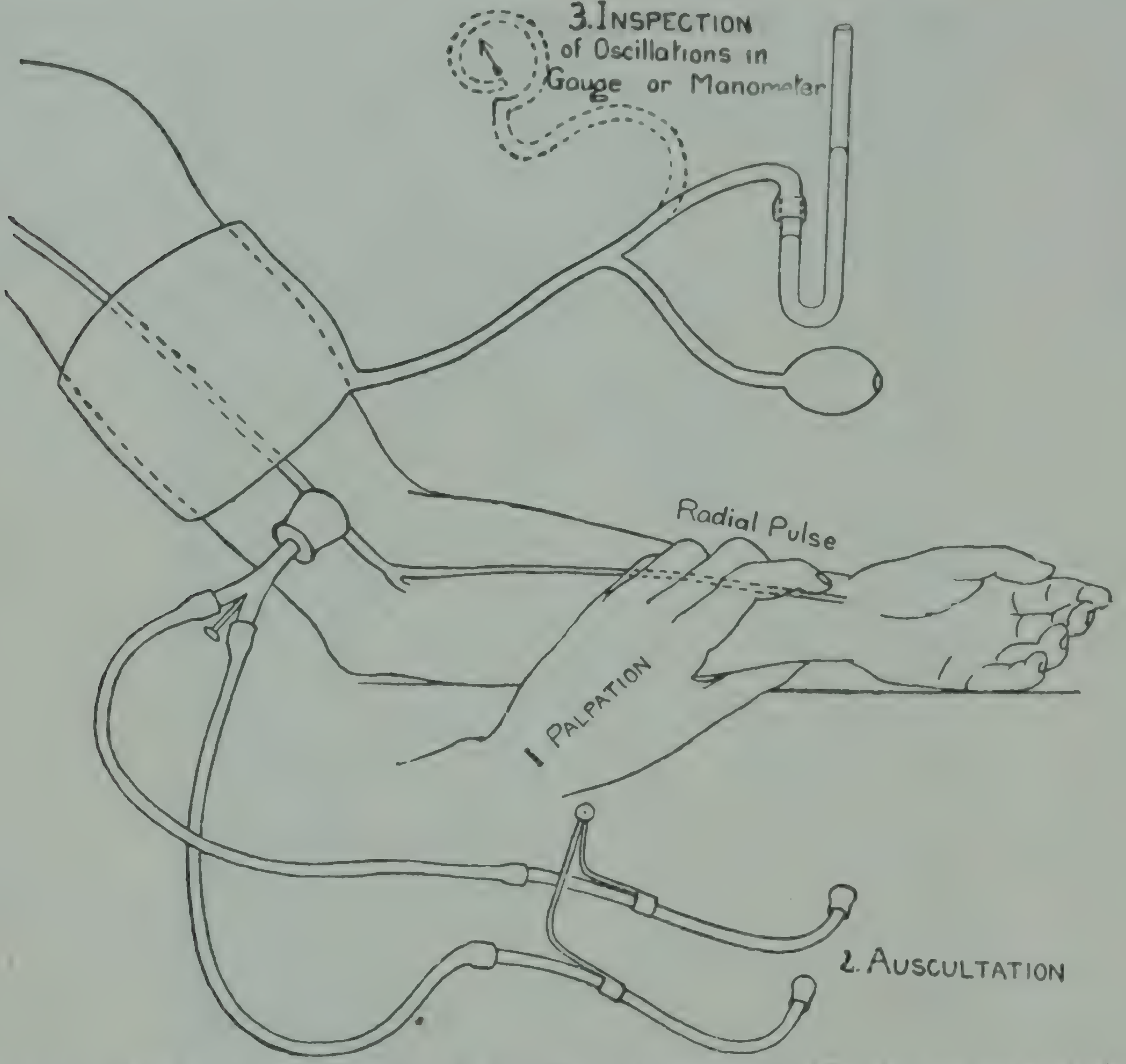


Fig. 411. Methods of using the Sphygmomanometer. (Outlet valve not shown.) (Harris, "Experimental Physiology.")

The principle of this method will be made clear by reference to the diagram (Fig. 412). If we imagine A as a segment of the brachial artery passing through the tissues which are surrounded by the rubber bag, we see that so long as the pressure in the interior of the artery is greater than that exerted by the tissues on the exterior, the artery will be patent and the pulse can pass through. If, however, the pressure in the tissues becomes greater than the maximum pressure inside the artery, the segment of artery will collapse (as in B), thus stopping the transmission of the pulse wave. If we exclude the elasticity of the tissues themselves, we may take the pressure in the bag as representing the pressure in the tissue fluids surrounding the artery, so that the pulse-obliterating pressure in the bag will correspond to the systolic pressure in the artery.†

By a slight modification of the apparatus it is also possible to determine the diastolic pressure. For this purpose the rubber bag is connected with a gauge of small inertia, giving a true representation of the actual changes of pressure. When the pressure in the bag just exceeds the diastolic pressure, the artery will be collapsed in diastole, and will then dilate almost to the utmost with the systolic rise of pressure. If we are taking a record of the pressure changes in the bag in this way, the pulse waves as recorded by the manometer will slowly increase in size as the pressure in the bag is gradually raised from zero. At one point the waves rapidly reach a maximum, and the pressure in the bag at this point is taken as the diastolic pressure. As the pressure is still further raised, the excursions of the manometer diminish again, first slowly and then rapidly, and the point of rapid diminution corresponds to the systolic pressure. Above this point the manometer still shows small oscillations, due to the impact of the stump of the occluded artery on the upper border of the bag.

---

* Riva-Rocci, *Gazz. med. di. Torino*, 1896, **51**, 52; Hill and Barnard, *B.M.J.*, 1897, ii., 904.

† Mummery. *J. Physiol.*, 1905, **32**, Pr. xxiii.

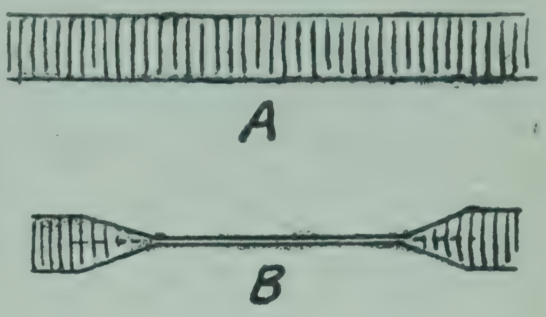


FIG. 412.

THE AUSCULTATORY METHOD. The most convenient method of finding systolic and diastolic pressures in the artery is to apply the bell of a stethoscope over the brachial artery a little distance below the cuff of the sphygmomanometer. When the cuff is rapidly inflated to a point beyond the extinction of the pulse, and then gradually deflated, a series of sounds ("Korotkov sounds") are heard, divided by Wiggers into five periods:

(1) As the pressure falls there is a sudden appearance of a clear sound, as the first flow of blood occurs through the compressed artery. This corresponds to the *systolic pressure.*
(2) The sound becomes more continuous, like a murmur.
(3) The sound becomes progressively louder.
(4) The sounds suddenly become muffled; this is the *diastolic pressure.*
(5) The sounds disappear.

The arterial pressure varies over a wide range in normal subjects, and over a narrower range from time to time in any given healthy subject. Values often quoted for adults are 120 mm. Hg systolic and 75 mm. Hg diastolic, usually written 120/75, but the errors of estimation are considerable, and it would be incorrect to accept these values as fixed standards of normality.* Apart from the incidence of disease, there is a tendency for arterial pressure to increase with age, *e.g.* at six months about 90/60, at four years about 100/65, at sixteen about 120/62, but different observers usually get different actual figures. A survey of systolic blood pressures in 7,216 apparently normal persons between sixteen and forty years gave the following results (Alvarez †):—

	Age.	
	18 years.	16–40 years.
Number of cases	1,216	6,000
Mean systolic pressure . . .	130 mm. Hg.	128·9 mm. Hg.
Standard deviation	13·4 mm.	13·5 mm.

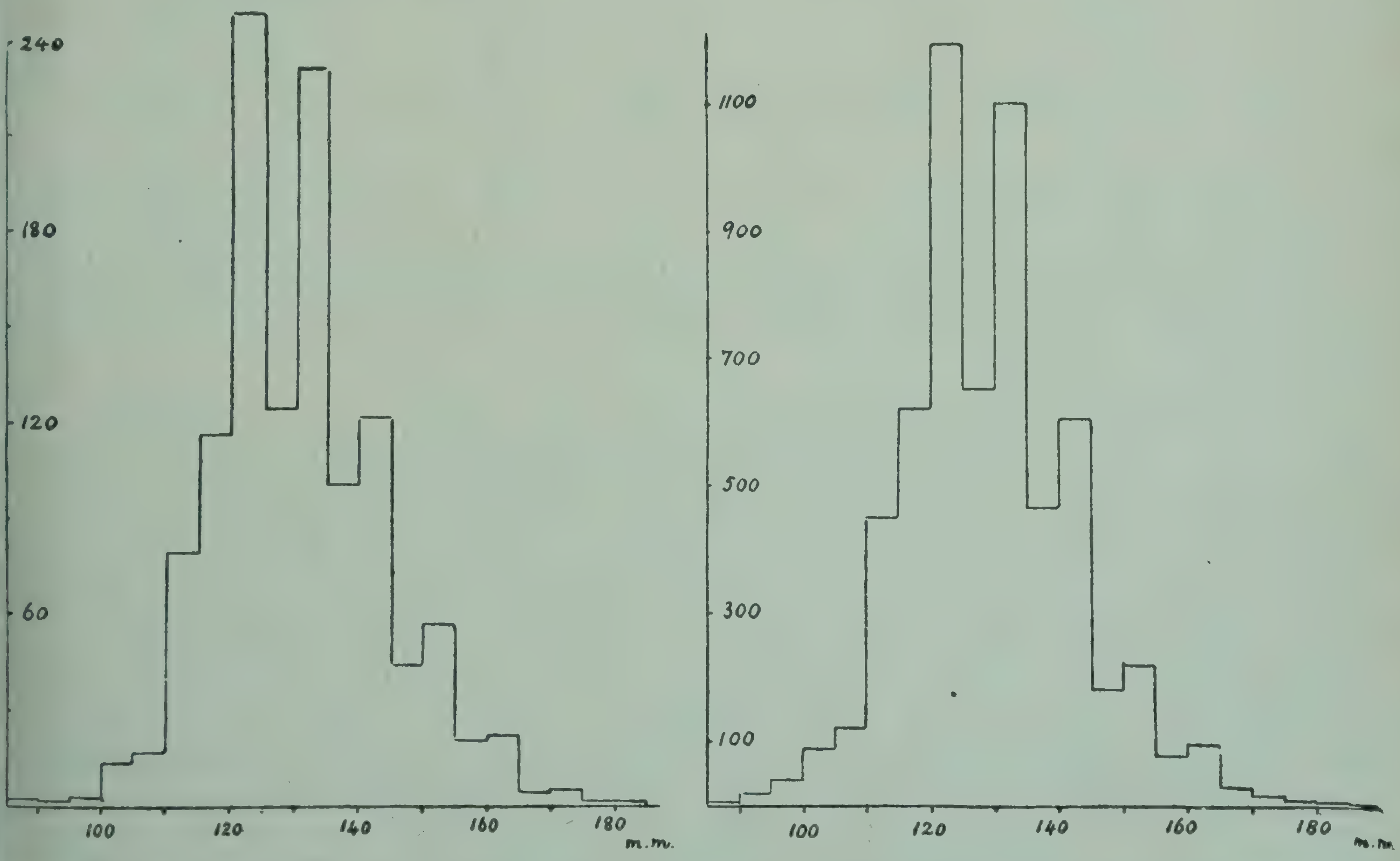


FIG. 413. Showing Frequency-distribution of the Normal Systolic Blood Pressure. (From Alvarez's tables.) Blood pressure as abscissa, number of persons as ordinate. On the left is shown the distribution of figures for 1,216 persons of 18 years of age. On the right figures for 6,000 persons from 16 to 40 years of age. (BURN. *Proc. Roy. Soc. Med.*, 1934, **28**, 15.)

* BURN. *Proc. Roy. Soc. Med.*, 1934, **28**, 5. TREADGOLD. *Lancet*, 8th April, 1933.
† ALVAREZ. *Arch. Int. Med.*, 1923, **32**, 17.

Unless we are to suppose that a proportion of these people were in an early stage of disease which at the time had produced no other effects, the figures mean that one-third of all males between the ages of sixteen and forty years of age have normal systolic pressures less than 116·5 mm. Hg or greater than 143·5 mm. Hg, *i.e.* one-sixth will have pressures normally over 143·5 mm. Hg. One man in about every forty will have a normal pressure greater than 157 mm. Hg, or less than 103 mm. Hg. The figures actually ranged from 85 mm. to 180 mm., in a distribution shown in the form of a frequency diagram in Fig. 413.

Venous Pressure. The pressure may be approximately estimated in the superficial veins in man by altering the elevation of the part containing the vein so that the vein just collapses: the horizontal distance of this level from the heart, or from some agreed surface marking, such as the manubrio-sternal junction, is then measured. This method of course only gives minimal pressures, the pressure in the feet when standing, for instance, being greater than when lying with the leg in such a position that the veins just collapse.

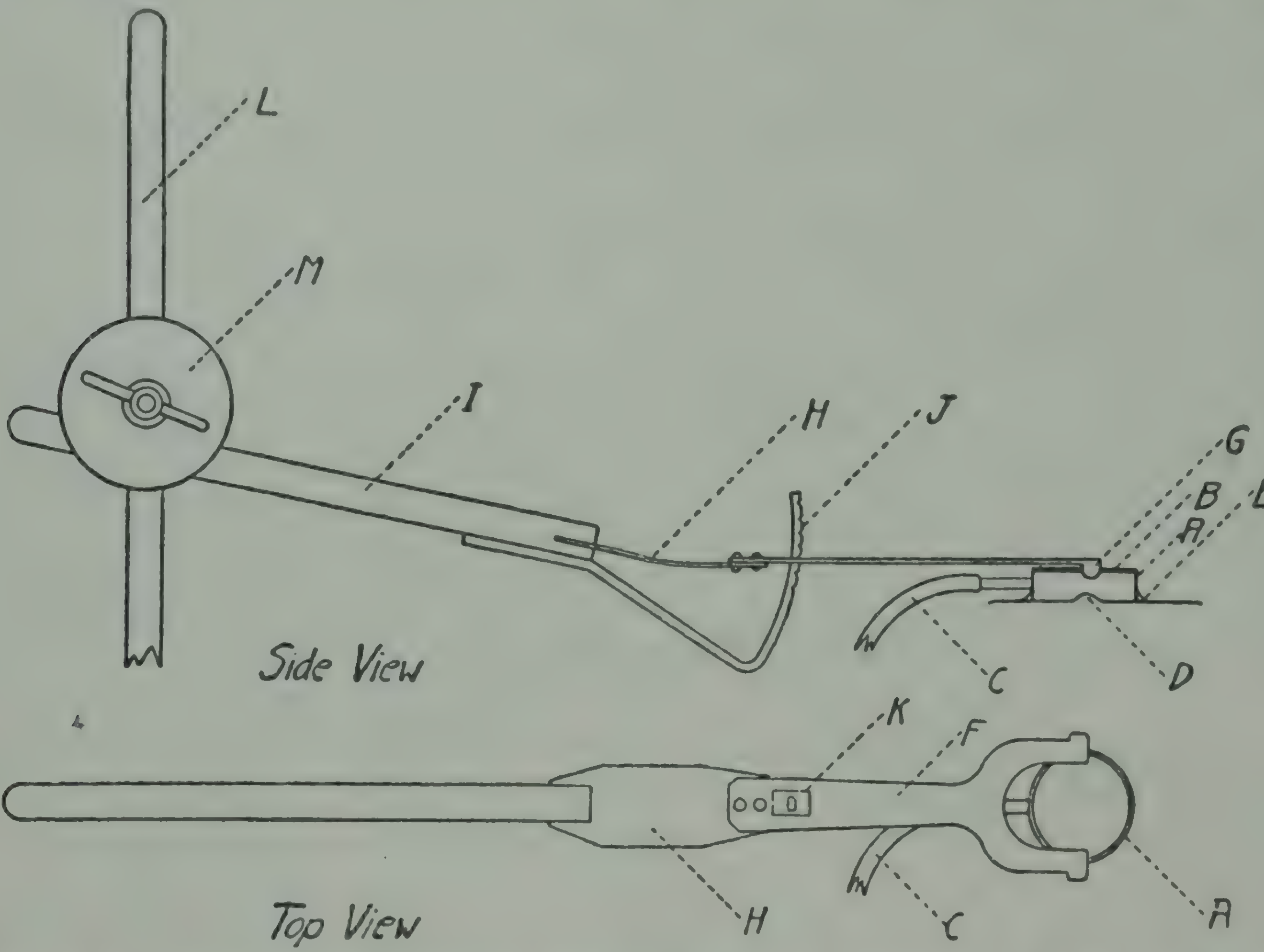


FIG. 414. Capsule for Measurement of Venous Pressures in Man. (KROGH TURNER and LANDIS. *J. Clin. Inv.*, 1932, **11**, 357.)

In order to determine the venous pressure directly the simplest way is to pass a hollow needle into the vein, and connect it with a citrate manometer by a connecting tube entirely filled with fluid. The level of the citrate solution is then referred to that of the manubrio-sternal junction; records of variations can be taken by connecting the open limb of the manometer to a tambour.*

Another method depends on the use of a principle introduced by v. Recklinghausen.† One of the modifications of this method is by Krogh, Turner and Landis ‡ (Fig. 414). A celluloid capsule A, with a celluloid top B, and a notch D on each side for the vein, is cemented to the skin over the desired vein, and held in position by a clamp having a steel spring H by which suitable pressure can be put on the capsule. A side tube C leads to a pressure bulb and water manometer, by which the air pressure inside the capsule can be altered to a known extent. The air pressure necessary to just collapse the vein is found, and is taken as equal to the venous pressure.

In the measurement of venous pressures it is customary to refer the hydrostatic level to that of the heart, but this is a purely arbitrary point and necessarily vague. It has in fact been shown that in the dead animal the reference point is caudal to the heart and is the same whether in the head-up or in the head-down position: in the living animal there are two reference points owing to the fact that the column of venous blood is interrupted by the heart; one point for the head section is above the heart and one for the tail section is below the heart.§ Owing to the uncertainty of the position of these reference points it is a common practice to measure the venous pressure levels from some definite anatomical point, *e.g.* the manubrio-sternal junction.

---

* DOUPE, KRYNAUW and SNODGRASS. *J. Physiol.*, 1938, **92**, 383.

† v. RECKLINGHAUSEN. *Arch. f. exp. Path. Pharm.*, 1906, **55**, 463. HOOKER. *Amer. J. Physiol.*, 1916, **40**, 43.

‡ KROGH, TURNER and LANDIS. *J. Clin. Invest.*, 1932, **11**, 357.

§ CLARK, HOOKER and WEED. *Amer. J. Physiol.*, 1934, **109**, 166.

For Reference

Eyster (1926). Venous Pressure and its Clinical Applications. *Physiol. Rev.*, **6**, 281.
Franklin (1928). The Physiology and Pharmacology of the Veins. *Physiol. Rev.*, **8**, 346.
MacWilliam (1925). Blood Pressures in Man under Normal and Pathological Conditions. *Physiol. Rev.*, **5**, 303.

ELEMENTARY HYDRODYNAMICS OF THE CIRCULATION

The blood flows through a branched closed system of tubes of elastic properties and of extremely variable diameters; moreover it is not a homogeneous liquid. Also, it is pumped intermittently into the arteries. The physical principles involved are consequently of great complexity and are for the most part as yet incapable of strict mathematical treatment. Nevertheless, certain elementary principles of hydrodynamics help towards a partial elucidation of the problems involved.

The cause of the pressure gradients in the different parts of the vascular system will be rendered clearer by a study of the flow of fluid through a tube of uniform bore (Fig. 415) from a reservoir, the level in which is maintained constant, and which therefore exerts a constant hydrostatic pressure. It really amounts to the statement that water flows downhill. If the tube AG be connected with the reservoir R, fluid will flow out from G under the influence of hydrostatic pressure. The pressure on the fluid at each part of the tube can be measured by attaching at a series of points, *e.g.* at B, C, D, E, F—vertical tubes in which the fluid will rise to a height corresponding to the lateral pressure existing at these several points. If AG is of uniform bore, it will be found that the heights of the fluid in the tubes show a continuous descent, so that the line joining the tops of the fluid in the various tubes is a straight one. The movement of the fluid from D to E can be regarded as due to the difference of the pressure between D and E, *i.e.* P_4—P_5. It will be noticed in the diagram that the straight line joining the tops of the fluid does not strike the surface of the fluid in R, but falls a little below it. Of the total potential energy represented by the pressure H in R, the large portion h' is employed in overcoming the resistance to the flow in the tube AG, while the remainder h represents the force which gives to the fluid as it leaves the reservoir at A a certain velocity. The potential energy of the flowing liquid diminishes as it moves along, as is shown by the fall of lateral pressure, but the kinetic energy is constant throughout in such a uniform tube. If the flow of fluid be diminished by partially clamping the end at G, or by increasing the length of FG, the rate of fall of the pressures will be diminished. This is because in such a type of flow the velocity varies inversely as the frictional resistance, and *vice versâ*. A smaller portion (h) of the hydrostatic pressure is now used to impart velocity to the fluid, and a larger portion, h', to overcome resistance. The same effect will be produced either by raising the level of G or by lowering the level of the reservoir and so the pressure at A. This is because the velocity, and hence the frictional resistance, depends on the pressure difference.

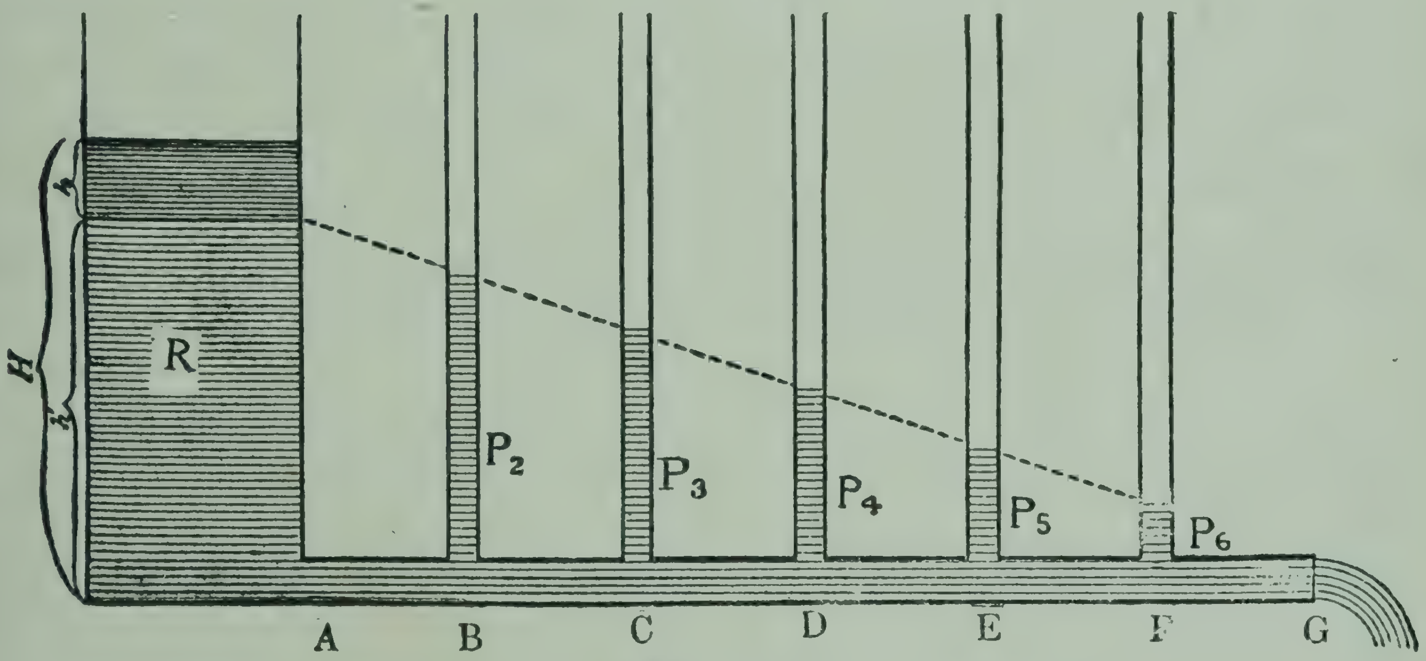


Fig. 415.

The difference of pressure between any two points, *e.g.* between D and E, may be regarded as that pressure which is necessary to maintain a certain volume-flow of the fluid against the resistance offered by the friction of the fluid in contact with the walls of the tube. This friction, and therefore the resistance to the flow, can be increased by diminishing the diameter of the tube, when a larger difference of pressure will be necessary in order to maintain the same flow. This can be shown by introducing a resistance between D and E by partially clamping the tube at this point (Fig. 416). The uniform fall of pressures in the vertical tubes is at once abolished. Between A and D there is a continuous, but slower, fall,

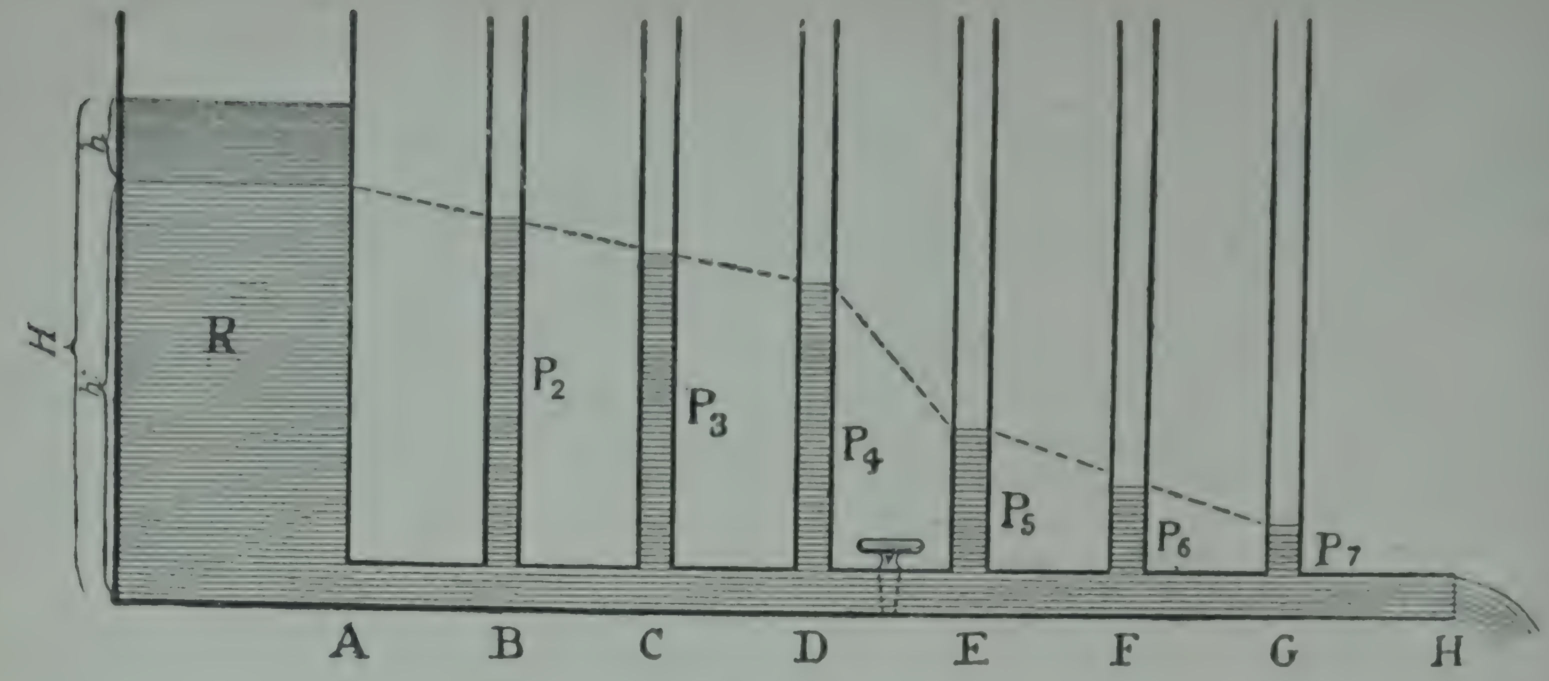


FIG. 416.

which is succeeded by a steep fall (P_4—P_5) between D and E, and this again by a gradual fall between E and G. Actually the volume of fluid flowing in unit time from a constant pressure source steadily along a straight tube varies as the fourth power of the radius of the tube and as the pressure difference between its ends, and inversely as the length of the tube and the viscosity of the liquid (η). This is Poiseuille's formula : Vol. $= \frac{\pi P r^4}{8 \eta l}$.

or resistance to the volume flow, $R = \frac{8 \eta l}{\pi r^4}$.

In the scheme shown, the reduction of the diameter of the tube at one point necessitates a greater difference in pressure between D and E than between the other points, in order to drive equal volumes of fluid through each section of the tube ; the total flow is also reduced for the reasons given above. The mean velocity of movement of the liquid is $\frac{\text{volume flow}}{\pi r^2}$ and hence, from Poiseuille's formula Vel. $= \frac{P r^2}{8 \eta l}$. In any such system of tubes, therefore, through which fluid is flowing, the fall of pressure between any two points will be $= \frac{\text{Vel} \times 8 \eta l}{r^2}$, *i.e.*, will be proportional to the product of the velocity and the resistance to the flow $\left(= \frac{8 \eta l}{r^2}\right)$ between these two points. The velocity on the other hand, $\left(= \frac{P r^2}{8 \eta l}\right)$, like the volume flow, will vary directly as the difference of pressures, and inversely as the resistance between the two points. These relations may be expressed by the simple formulæ

$$P \propto VR,\ R = \frac{P}{V} \text{ and } V \propto \frac{P}{R}.$$

The scheme we have just considered is only capable of limited applications to the problems of the circulation. A nearer approximation could be got by studying the problem, not of a flow from a constant pressure source, but that of a constant *volume* of fluid in unit time, as from a rotary pump, along the tube of uniform bore (Fig. 417). This would more nearly represent the conditions in the circulation, except for the fact that in the latter the flow is intermittent. Again we have a gradient of pressure, as shown by the vertical tubes A to F ; the gradient will be steeper the narrower the horizontal tube and the greater the volume of fluid delivered in unit time by the pump ; *i.e.* its steepness will depend on the energy needed to overcome the frictional resistance in the tube. At the free opening G, the fluid will have no lateral pressure, but will have a certain kinetic energy, dependent on its velocity, *i.e.* on the difference between the potential energy at the source and the frictional loss in the tube. If now an increase of resistance be put in at X, as by partly closing a tap, the effect will be of the same nature, but much more marked than with the former scheme, because now there can be no question of the resistance causing a diminution of the flow, which by definition, is constant. The effect will be a considerable rise of pressure in the near side and an equally pronounced fall on the far side of the resistance (Fig. 417).

Next, let us consider the effect, on the same scheme, of altering the volume of fluid flowing through the horizontal tube. It will be found that the greater the volume delivered in unit time the higher the lateral pressure near its entry will be, and the steeper the gradient (Fig. 417).

We have supposed in all these examples that the movement of the fluid is "stream-lined." In this, the simplest type of fluid motion, the liquid moves in a straight line as concentric fluid tubes ; the layer in contact with the walls of the tube is stationary, the next layer glides over it slowly, the next one over that, and so on until, at the axis of the tube, we have the maximal velocity. In such a case the resistance, which is due to the friction of the layers over one another, varies directly as the mean velocity. In normal conditions the flow of blood in the arterial system is stream-lined. When, instead of being stream-lined the flow is whirling or turbulent, which happens above a certain critical velocity, even in straight tubes, the resistance varies approximately as the square of the mean velocity.* A formula used by engineers for such cases is $R = kSv^2$ where R is resistance, S is the surface area of the wall, and v the velocity. The lower the viscosity of the fluid, and the narrower the tube, the more readily does turbulent motion ensue, and hence the greater becomes the resistance. Murmurs are produced when the flow in large vessels is turbulent.

We may now apply some of these principles to the vascular system. In the living vascular system, the largest difference of pressure exists between the arteries on the one side and the small veins on the other, a great fall occurring between the small arteries and the capillaries themselves. This distribution of pressure points to the chief gradient in the vascular system as being situated in the arterioles. The resistance presented by these vessels is due to their small calibre, since they are maintained in a state of tonic contraction.

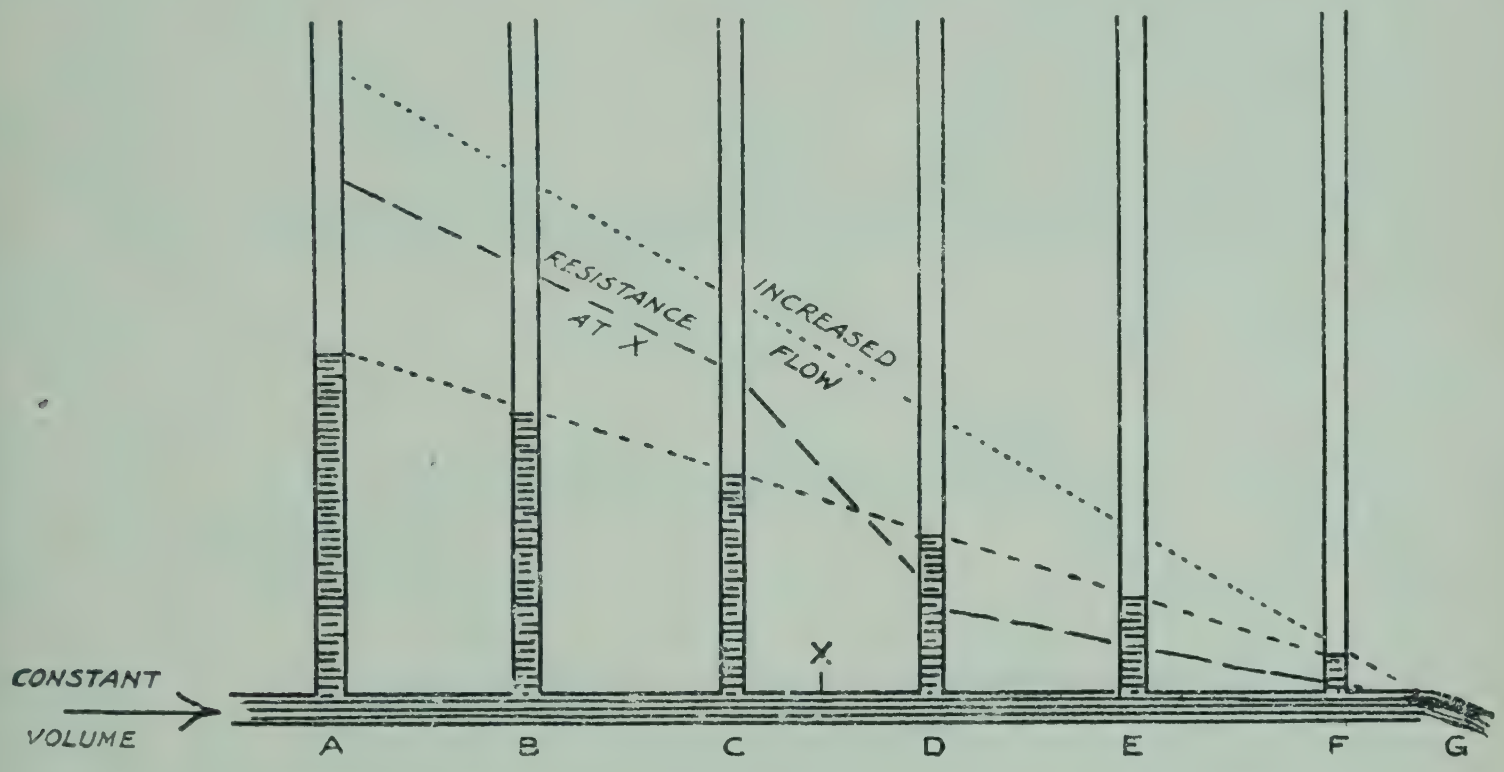


FIG. 417.

The total bed of the stream in the region of the arterioles, while greater than that of the arteries, is considerably less than that of the rich meshwork of capillaries, while the difference between the diameters of arterioles and dilated capillaries is not very great. On this account the velocity of the blood in the arterioles is very much greater than that obtaining in the capillaries, and since friction, and therefore the resistance, probably varies as the square of the velocity, the resistance to the flow of blood through the arterioles must often be greater than that presented by the capillaries. When the capillaries constrict, however, they also add largely to the peripheral resistance. The large part taken by the arterioles in determining the difference of pressure between the arteries and veins is shown by the fact that this difference can be diminished to one-half by any means which causes a general dilatation of the arterioles, as for example, destruction of the vasomotor centre.

As regards the occurrence of turbulent motion, Muller † found that in an elastic tube only 7·5 mm. in diameter the flow began to be turbulent at a velocity of about 50 cm. per second. Hence, quite apart from turbulence due to branching, the flow in the larger arteries could be a turbulent one.

The mean arterial pressure, *i.e.* the average pressure within the arteries at any given time, is thus dependent on :

(1) The resistance to the outflow of blood from the arterial system, *i.e.* on the state of constriction of the arterioles and, to a smaller extent, of the capillaries.

(2) The output of the heart per unit of time.

If one factor is known, the other can be computed, *e.g.* the total peripheral resistance $= \frac{P}{V} = \frac{\text{mean pressure}}{\text{cardiac output}}$.

The Effect of Branching. When a tube gives off a branch, the pressure exerted in the branch by a fluid *flowing along* the tube will depend upon the angle at which the branch leaves the main vessel. If, as at A, Fig. 418, the branch leaves at right angles, then the hydrostatic pressure in the branch is equal to the lateral pressure on the walls of the main vessel. If the branch inclines at an acute angle to the direction of flow, as at B, then owing to the kinetic energy of the moving column of fluid, the pressure in B will be greater than the lateral pressure in the main stem, while if at an obtuse angle, as at C, it will be less. This is the principle of what is known as the *piezometer*, and it follows that if B and C are connected together, as shown in the lower diagram of Fig. 418, there will be a subsidiary (so-called piezometric) circulation of fluid along the side system in the direction indicated by the arrows. This is in fact what happens very extensively in many parts of the circulatory system, particularly in the smaller arterial anastomotic circulations, *e.g.* from arch of aorta to internal mammary artery, then *viâ* ascending aortic intercostal arteries to the descending aorta.*

---

* WIEN-HARMS. *H'dbch. d. exp. Physik.*, 1931, IV. (i.), 293.

† *Ztschr. f. d. ges. Exp. Med.*, 1924, **39**, 210.

Most of the arteries give off their branches at an acute angle, so that the end-on pressure in the branch is greater than it would otherwise be.

The Effect of Viscosity. For reasons already explained (p. 501), the apparent viscosity of blood, as measured by the rate at which it flows along a capillary tube under given conditions, is variable. Direct experiment shows that in passing through the capillaries of the vascular system, blood behaves as though it had about twice the viscosity of water, although glass capillary viscosimeters usually give a viscosity at least double this. The reason for the low viscosity in the circulation is due to the small diameter of the vessels in which the main resistance is met. In such narrow vessels the corpuscles tend to move like a rod in the middle of the stream, leaving the plasma at the periphery to encounter the frictional resistance of the wall.†

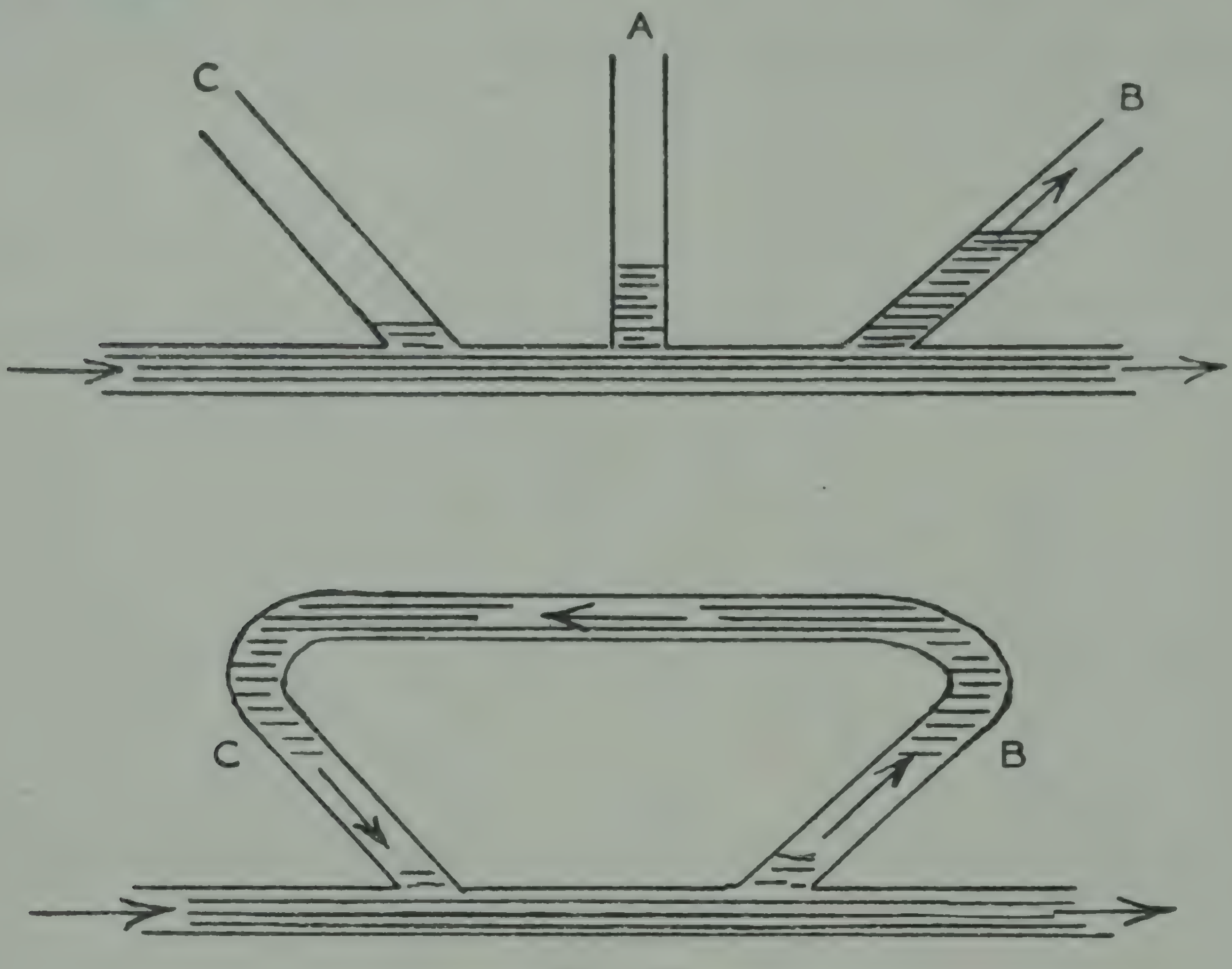


Fig. 418.

Increase in the corpuscular content increases the viscosity, *i.e.* more pressure is required to maintain a given flow along a given capillary ; in other words, if the corpuscular content is increased, a given capillary tube will offer a greater resistance to the flow (Fig. 419).

THE EFFECT OF CAPACITY. So far, we have considered only the influence of changes of flow, pressure and resistance in a simple tube with a head of pressure at one end and a free outflow at the other. In the body, however, the vascular system is a closed circuit of branched elastic tubes presenting varying resistances to the flow of blood, and of varying diameters and distensibilities at different parts of their course. In this closed system is inserted a pump which drives the blood through the system. Since all the blood vessels are elastic and distensible, the capacity of the system must vary with the internal pressure to which the vessels are subjected. Moreover, the position of the different parts of the body must have an influence on the capacity of the system, since the dependent vessels will be distended, not only by the average pressure of the fluid throughout the system, but also by the hydrostatic pressure due to the weight of the column of fluid above them. The elasticity of the tubes is also a varying factor, and can be considerably altered by the contraction of the muscular coats of the vessels, or by pressure exerted on the vessels by the surrounding muscular and elastic structures.

---

* Swindle. *Amer. J. Physiol.*, 1930, **93**, 588 ; Berry and Daly, *Proc. Roy. Soc.*, 1931, **109B**, 333.

† Whittaker and Winton. *J. Physiol.*, 1933, **78**, 339.

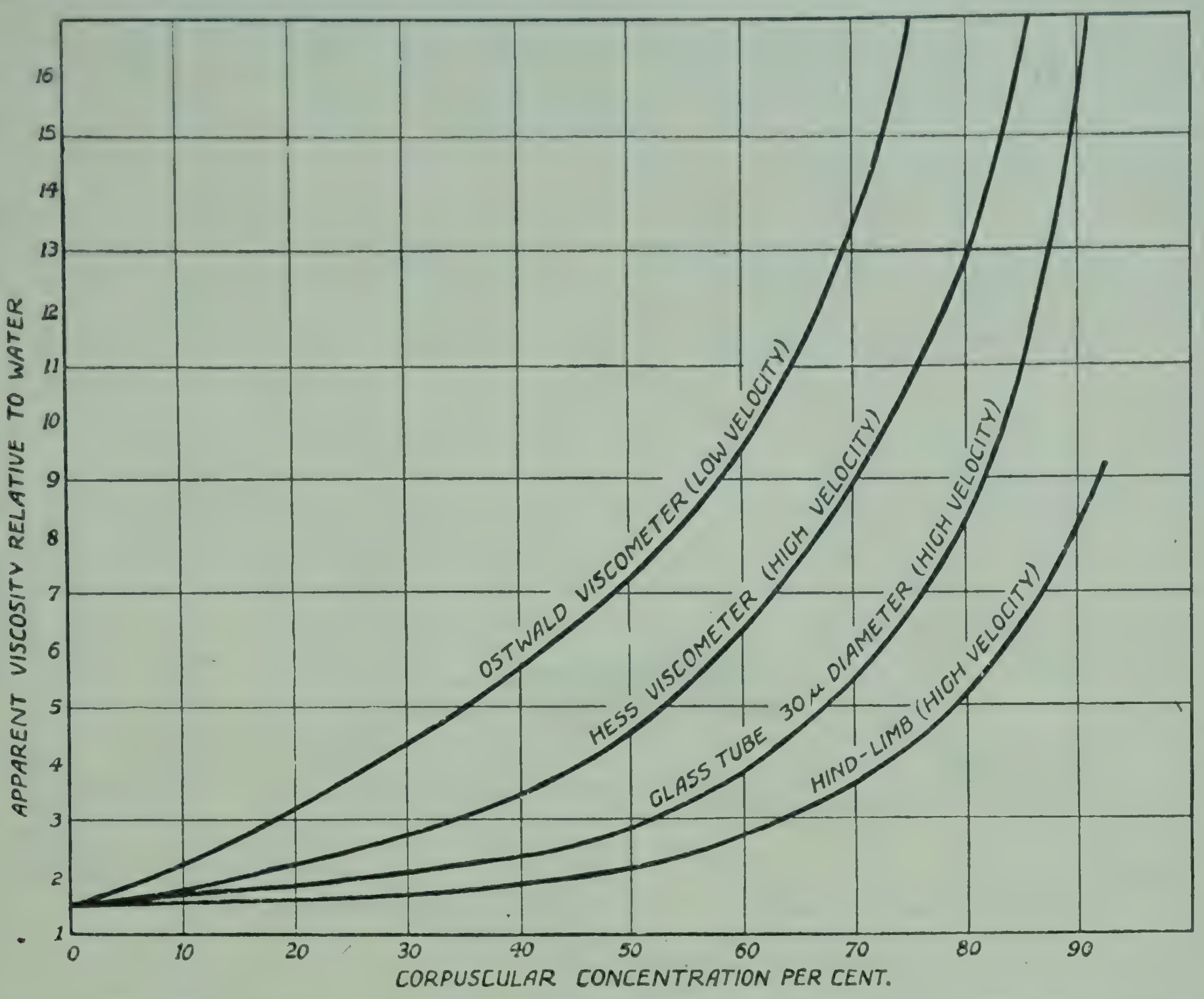


FIG. 419. The Influence of the Corpuscular Concentration (measured with a hæmatocrit) on the Apparent Viscosity of Blood.

Measurements were made on dog's blood, defibrinated, with four different types of viscometer: (1) the Ostwald (low velocity), (2) the Hess (high velocity), (3) a glass tube of very small diameter, and (4) the dog's hind limb. If the glass tube used had had an even smaller diameter, the apparent viscosities would have been still lower. (After WHITTAKER and WINTON and L. E. BAYLISS.)

Circulation Schema. It will simplify the discussion of the main factors of the circulation if, for the present, we neglect the variable factors and see what would take place in such a closed system of elastic tubes all situated on one horizontal plane. Such a system is represented in the diagram (Fig. 420), and a working model of it in Fig. 421.

The heart H is interpolated at one part of the circuit, while the free flow of the fluid from B to D is impeded by the presence of a peripheral resistance at C. Such a system would have a definite capacity at zero internal pressure, but a greater one under a positive pressure. We will assume that the pressure throughout the system when at rest is 10 mm. Hg., *i.e.* the elastic tubes are all slightly distended. If the heart H now begins to contract, it will pump fluid from E into A. The pressure in E will fall, while that in A will rise, the resistance at C preventing the free escape of fluid from B to D, and so causing the fluid to pile up in A.

If the properties of the tubes were uniform throughout the system, the rise of pressure in A would approximate very nearly to the fall of pressure in E. In the vascular system the veins are, however, much more easily distended or collapsed than the arteries

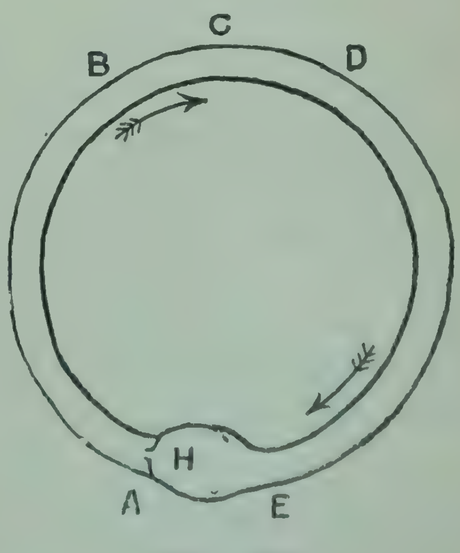


FIG. 420.

(Figs. 421 and 422). The vein collapses when there is no distending force in its interior. The slightest rise of pressure causes a considerable increase in its capacity, which rises rapidly with increasing pressure up to about 20 mm. Hg. Whereas the artery needs a pressure of about 100 mm. Hg to distend it, and continues to stretch, but more slowly, at higher pressures, the vein is distended almost to its maximum at about 10 mm. Hg. Therefore, the tubes at E are made of thin-walled rubber, which will be considerably distended under a pressure of 10 mm. Hg, that has practically no influence on the thicker-walled arterial tube A.

A small amount of fluid taken from E would thus cause very little fall of pressure on this side. A considerable force will be necessary to send this fluid into the more resistant arterial tube, so that on pumping enough fluid from E to A to lower the pressure in E, say 5 mm., the pressure in A has to be raised to say, 100 mm. Hg in order to distend the arteries to such an extent that they will accommodate the extra fluid.

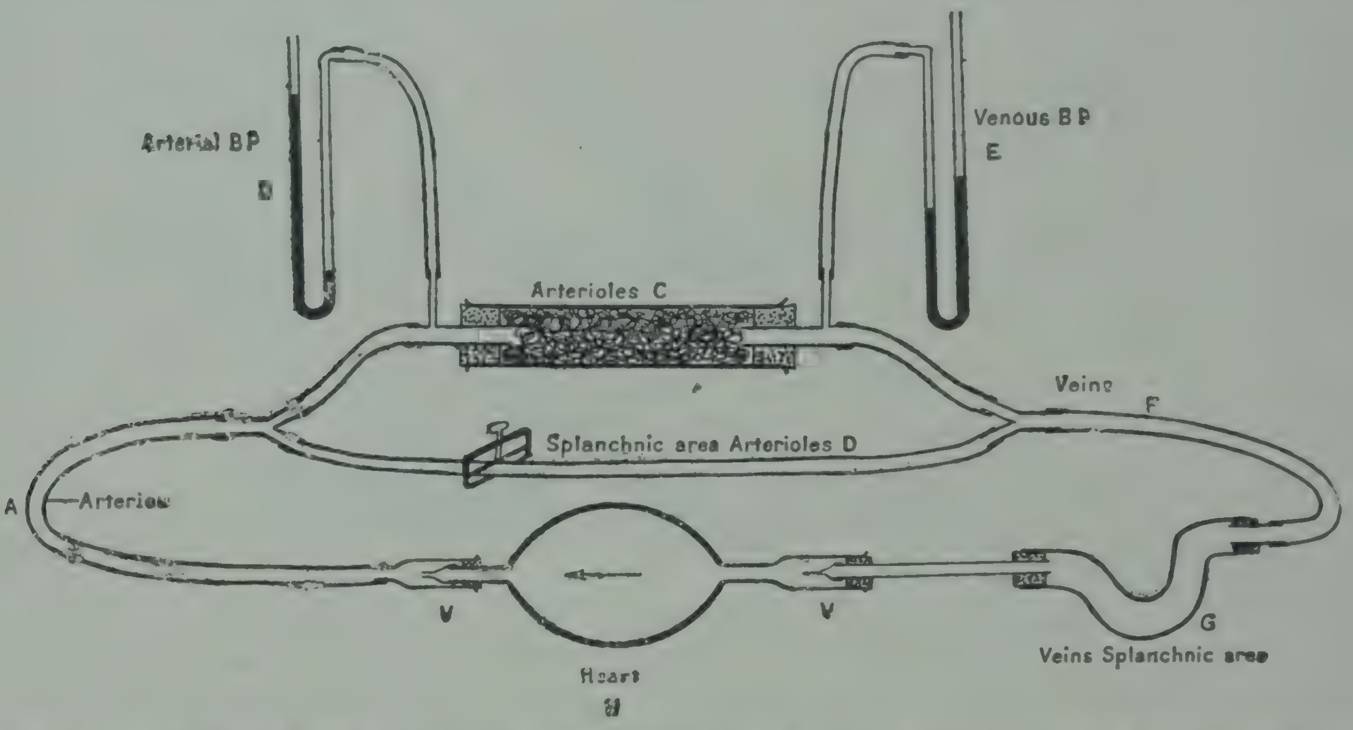


FIG. 421. Artificial Schema to Demonstrate the Main Features of the Circulation.

The heart is an enema syringe with valves at V and V. The artery is a thick-walled rubber tube. On the venous side is placed a length of wide thin-walled tubing, to represent the large thin-walled distensible veins. The arterioles and capillaries (peripheral resistance) are represented by wide glass tubes packed with sponges. By opening the clamp on the tube D ("splanchnic area arterioles") the peripheral resistance can be removed, and a free passage of fluid allowed from arterial to venous side.

In such a system, when the heart contracts, it takes fluid from the veins to the arteries, and piles up the pressure on the arterial side until it is sufficient to cause fluid to flow through the peripheral resistance into the veins at the same rate as it is taken by the heart from the veins. This rise of pressure in the arteries may be many times greater than the fall of pressure in the veins. If more fluid is injected into the system, the whole system will be more distended and the *mean systemic pressure will rise.* Then, according to the rate of output by the heart, the arterial pressure may be less than, equal to, or greater than the pressure attained before the introduction of fluid. Supposing it to be the same as before, the increased amount of fluid must be accommodated somewhere, so that the venous side must be fuller. In the same way, the withdrawal of a certain amount of fluid may lower the mean systemic pressure, and it would still be possible for the pump to maintain an arterial pressure equal to that produced when the mean systemic pressure was 10 mm. Hg, but to obtain this effect the rate of its output would have to be increased, and the veins would be more empty than they were previously. The maintenance of a constant arterial pressure with varying amount of fluid in the system can therefore be accomplished either by alterations in the output of the heart or by alterations in the peripheral resistance, and therefore in the ease with which the blood is allowed to pass from the arterial to the venous side.

Alterations of the capacity of the system will have the opposite effect to alterations of its contents. Thus diminution in the volume of veins, such as might be caused in the living body by pressure on the veins from without, will drive the fluid into other parts of the system and therefore raise the mean systemic pressure. This rise of pressure may be confined to the arteries by increased action of the heart, or it may be confined to the veins by diminished action of the heart or decreased constriction of the arterioles forming the peripheral resistance.

Similar change in capacity may be brought about by alterations in the hydrostatic pressure in the vessels. If, in the model illustrated (Fig. 421), we allow the thin-walled vein to hang over the edge of the table, the pressure of the column of fluid within it causes it to distend and to accommodate more fluid, and this increased capacity might be so great that the pressure in the section of the vein near the heart might fall so low that the heart would receive no blood. The whole arterial system might in this way drain under the influence of gravity into the large veins.

All the conditions in our artificial schema have their exact analogue in the living body. The determination of the mean systemic pressure in the living body is difficult to carry out with accuracy. Direct measurement shows that it is not more than 10 mm. Hg. Under the influence of gravity, the pressure would be easily reduced to zero, or, in the upper parts of the body, even to a pressure below that of the air, but for the nervous reactive mechanism which we shall consider later on, and the whole of the blood would accumulate in the abdomen and lower parts of the body, and the circulation would come to a standstill.

Widespread or universal dilatation of the capillaries, as may occur under the influence of histamine, or in some forms of "shock", may so increase the total capacity of the vascular system that the blood remains in the capillaries, and an ever decreasing amount of blood is available to pass on into the heart. The heart thus has a diminished output, and is unable to maintain the normal arterial pressure in the arterial system.

The *mean pressure* may be altered in the vascular system in the following ways:

(1) Alteration of capacity of the total system, either by contraction of the walls of the vessels or by pressure on them from without.

(2) Alteration of the total volume of the circulating fluid.

The *distribution of pressure*, *i.e.* the relative pressure in the arteries and veins, will be determined by

(3) Alteration in the output of the heart. This may be augmented by greater amplitude of ventricular contraction increasing the output per beat, or by greater frequency of beat. Generally both means are employed.

(4) Alteration in the peripheral resistance and therefore in the ease with which the blood can pass from arterial to venous side.

In any change, either in arterial or venous pressure, at least two of these factors are involved. Nearly always such a change will involve as its immediate consequence some corresponding alteration in the heart beat, so that at least three factors will usually co-operate in the production of the rise or fall of blood pressure.

The Dependence of Arterial Pressure on Output of Heart. Arterial pressure is a resultant of the two effects:

(*a*) The rate at which blood enters the arterial system from the heart;

(*b*) The rate at which blood leaves the arterial system through the peripheral resistance.*

---

* For hydrodynamical treatment *v.* BROEMSER and RANKE. *Z. f. Biol.*, 1930, **90**, 467; WEZLER. *Ibid.*, 1938, **98**, 438.

It is evident that the arterial pressure will be altered by varying either of the two factors—output of the heart or peripheral resistance. The cardiac output will depend on the stroke volume of the heart and on the pulse rate. The filling of the heart at the beginning of each beat is in its turn dependent on the pressure in the great veins. If the heart is beating with optimum rate and force, it will reduce the distension of the veins nearest the heart, and it is not possible for it to obtain more blood to put into the arterial side, however frequently it may beat. Increased frequency of heart beat would not, therefore, necessarily increase the total output of the heart. There will be an optimum frequency of the heart beat which will depend on the state of filling of the great veins. The fuller these are, the more the total output can be augmented by an increase of the heart rate. On the other hand, slowing the heart from its optimum rate will increase its stroke output, but diminish the minute output.

BLOOD FLOW

The Conversion of an Intermittent into a Constant Flow. By the time it has arrived in the veins, the flow of blood has been converted from a pulsatory into a continuous flow. This change is connected with the elastic nature of the arterial walls and will be more easily understood by a simple illustration. If a pump be connected with a rigid tube, every time that a certain amount is forced into the beginning of the tube an exactly equal quantity will be forced out at the other end. Increasing the peripheral resistance by partial closure of the end of the tube will not affect the intermittent character of the flow, but will merely increase the force necessary to expel the fluid. If, instead of a rigid tube, we employ a wide elastic tube, and the end be left open so that no resistance is offered to the outflow, the outflow will, as with a rigid tube, be intermittent. But now, if the end of the elastic tube be narrowed so as to increase the resistance to the outflow, there will be a marked difference. Each stroke of the pump forces the same amount of fluid into the tube as before, but owing to the peripheral resistance this does not now all escape at once ; instead, part of the force of the pump is spent in distending the walls of the tube, and part of the fluid that was forced in remains for a time in the tube. The distended tube tends to empty itself and force out the fluid which over-distends it before the next stroke of the pump occurs. So now the outflow may be divided into two parts, one part which is forced out by the immediate effect of the stroke of the pump, and another part which is forced out by the elastic recoil of the tube between the strokes. If the strokes be rapidly repeated before the tube has time to empty itself thoroughly, it will get more and more distended. Greater distension means stronger elastic reaction, and therefore a larger outflow of the fluid between the beats. This distension goes on increasing till the fluid forced out between the strokes by the elastic reaction of the wall of the tube is exactly equal to that entering at each stroke, and the flow thus becomes continuous.

The same thing occurs in the living body. A man's heart at each beat forces about 60 c.c. of blood into the already distended aorta. The largest arteries, such as the aortic arch, innominate, and subclavian arteries are highly distensible, and so the expelled blood is readily accommodated in them ; this wave of distension is transmitted with gradually decreasing force along the arteries and is what we feel as the pulse. After each heart beat the arteries tend to return to their original size, and drive the blood onwards. By the time the blood has reached the veins, all trace of the heart beat has disappeared.

VELOCITY OF THE BLOOD

The velocity of the blood as a whole will vary directly as the rate of the cardiac output, but, apart from changes in this, the speed of movement of the blood in different parts of the circulation will always vary greatly, for reasons already explained.

We may consider (1) the circulation time, *i.e.* the time taken for blood to traverse a part or the whole of the circulation; (2) the linear velocity of the blood as it passes a particular spot, and such velocity may be considered either as a mean velocity or as a velocity fluctuating from moment to moment; (3) the volume of blood flowing through a particular part of the circulatory system.

(1) **Circulation Time.** Measurements of the circulation time are made by introducing into the blood stream some easily recognised substance at a given point, and finding how long it takes for that substance to appear at any other given point in the circulatory system. If the tell-tale substance is injected into a vein, and the time taken for it to appear in the blood of the same vein distal to the point of injection, or in the corresponding vein on the opposite side, then we get the *total circulation time;* if the transit is from vein to lungs we get a measure of the time for the pulmonary circuit; if from vein to capillaries we get the time for flow from vein through pulmonary circuit, left heart and arteries to capillaries, *i.e.* total circulation time less the time for venous, return, etc.

Substances which can be used for injection are: (1) Potassium ferrocyanide, in animals; (2) radioactive material, detected at various points, without drawing blood, by the use of an ionisation chamber; (3) the dye fluorescin, for total time; (4) acetone or ether, detectable in the breath, for pulmonary circuit time; (5) histamine (0·001 mg. phosphate per kg.), which causes flushing of capillaries of the face, or saccharin, which gives a sweet taste when it reaches the tongue, for the vein-heart-capillary time. (6) A small dose of NaCN, which causes hyperpnœa when it reaches the carotid body.

Typical times found are:

Total circulation time (vein to vein)	12 to 28 secs. (av. 25)
Arm vein to face (histamine)	24 secs. (saccharin 11 secs.)
Arm to artery (radio-active method)	14 to 24 secs.
Vein to lungs (ether or acetone)	6 secs. (radio, 11 secs.)
Arm to heart (radio-active method)	2 to 14 secs.
Arm vein to carotid body (NaCN)	12 to 15 secs.

It will be evident that the time measured in all these methods is the shortest time taken by any of the blood to traverse the given part of the circulation, and hence the figures are variable unless exactly the same method is used. For different animals the circulation times are shorter the smaller the animal, as would be expected. The number of heart beats per circulation time, is, however, about the same for all species, viz. about thirty beats per total circulation time.

Factors which increase the cardiac output will tend to shorten the circulation time. Hence it is reduced in exercise, after adrenaline and in such diseases as hyperthyroidism, where the total metabolism is raised, or anæmia, in which the blood volume and viscosity are diminished. It is increased in myxœdema, because the basal metabolism is reduced, in polycythæmia, because the viscosity of the blood is increased, and in uncompensated disorders of the heart.

(2) **The Velocity of Flow at Different Points.** It is evident that an increased output of the heart in unit time will speed up the rate of flow of blood in all parts of the circulation. At any given rate of output, however, the velocity of the blood varies as it passes along the vessels. If a liquid be flowing steadily along a tube of unequal bore, as represented in Fig. 422, since the amount of fluid passing *a* during a given interval of time must be equal to the amount passing *b*—where the bed of the stream is wider—the velocity of the flow must be smaller at *b* than at *a*.

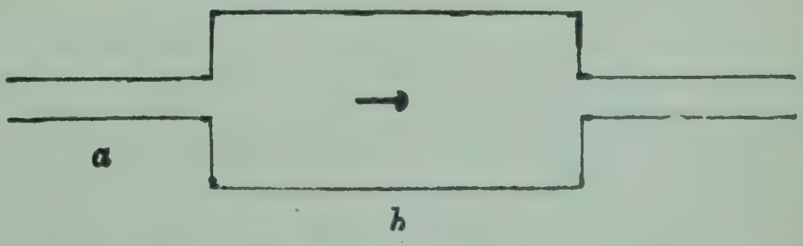


Fig. 422.

It is obvious that the velocity of flow at any two points in the same system will vary inversely as the cross-sectional areas of the whole bed at those points. Thus in a closed circuit (Fig. 420) with a steady flow from the arterial to the venous side, the amount of fluid passing A during a minute must be exactly equal to the amount passing through the peripheral resistance C. The total area at C is, under average conditions in the body, probably one thousand times that of the aorta at A, and we should expect therefore a proportionate slowing of the blood stream. As a matter of fact, while the mean velocity of the blood in the aorta of a large animal may be taken as about half a metre per second (varying at different parts of the cardiac cycle from 0·1 to 1·0 metre per second), the velocity of the blood in the capillaries is about half a millimetre per second. Moreover, since the total cross section of the big veins near the heart under a normal distending pressure is about twice that of the first part of the aorta, the velocity of the blood in the great veins is only about half of that found in the aorta.

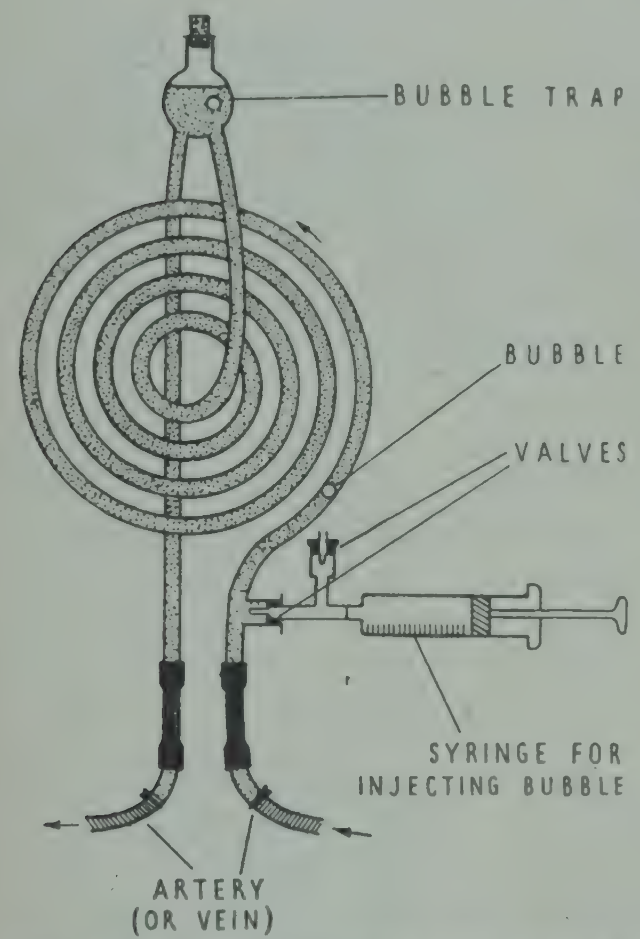


FIG. 423. Diagram of the bubble *Stromuhr*, as described by DUMKE and SCHMIDT. (*Amer. J. Physiol.*, 1943, **138**, 421).

The foregoing discussion must not be applied directly to the changes in the velocity in any part following on *local* alterations in the resistance. In this case the local changes are insufficient to affect the general arterial blood pressure, and the result of diminution of perpiheral resistance is to furnish a short cut from the arterial to the venous side for a portion only of the total output of the heart. Thus, dilatation of the vessels of the submaxillary gland, while not altering the general blood pressure, causes the blood flow through the gland to be increased six to eight times. The pressures and velocities in the *local* arterioles, capillaries and veins are all increased by this local vaso-dilatation. Conversely, constriction of the arterioles of any given part will diminish the velocity of the blood through this part and also the pressure in its capillaries.

The larger the area affected by the change in the peripheral resistance, the more difficult it is to predict *a priori* what will be the result on the velocity of the blood and on the circulation as a whole, or in the parts specially affected. In the living body a widespread dilatation of the arterioles increases the total capacity of the system. The arterial relaxation, therefore, not only gives rise to an easier outflow from arteries to veins, but may also cause a diminished filling of the veins, and therefore of the heart during diastole. The heart output is therefore also lessened, so that a final result of a dilatation of the arterioles may be a diminished, instead of an increased, velocity throughout the system.

METHODS OF MEASURING THE FLOW AND VELOCITY OF THE BLOOD. If we can measure the average volume of blood passing along a given vessel per minute, we can calculate the mean speed of its movement. In experimental animals, the mean flow can be measured by a *Stromuhr*, of which there are many types. A modern form is shown in Fig. 423. Its principle should be obvious. The passage of the heparinised blood along the spiral, which is inserted in the course of a blood vessel, is timed by measuring the time taken for a bubble of air injected by the syringe to pass from a mark near the syringe to a second mark beyond the spiral. The capacity of the tube between these two marks being known, a simple calculation will show the amount of blood passing along the vessel per minute.

Another form of stromuhr is the *thermostromuhr*.* The principle is that a spot

---

* REIN. *Z. f. Biol.*, 1929, **89**, 195. BARCROFT and LONGBRIDGE. *J. Physiol.*, 1928, **98**, 382.

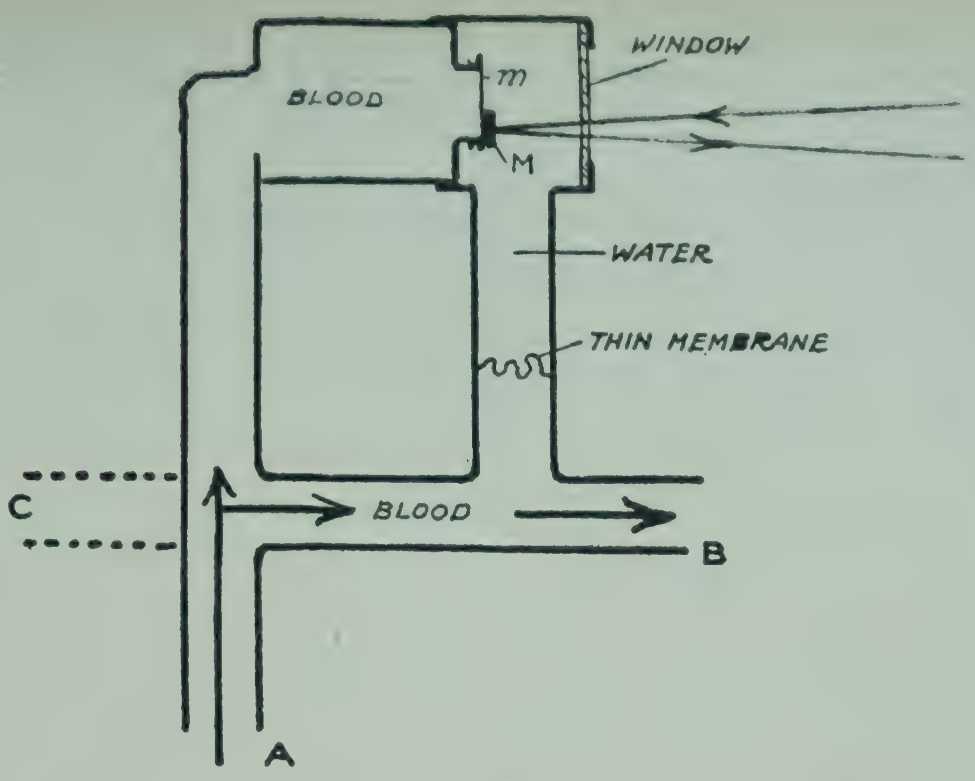


FIG. 424. Principle of the Fleisch Differential Stromuhr. The difference in pressure in the two branches is communicated to the thin membrane manometer, *m*, bearing a mirror M. The outer chamber is filled with water, so that light passing through the window can be reflected from the mirror, and on to a photographic surface, whereby the pressure differences can be graphically recorded. The thin *lax* membrane separates the water from the blood, but does not interfere with the pressure transmission.

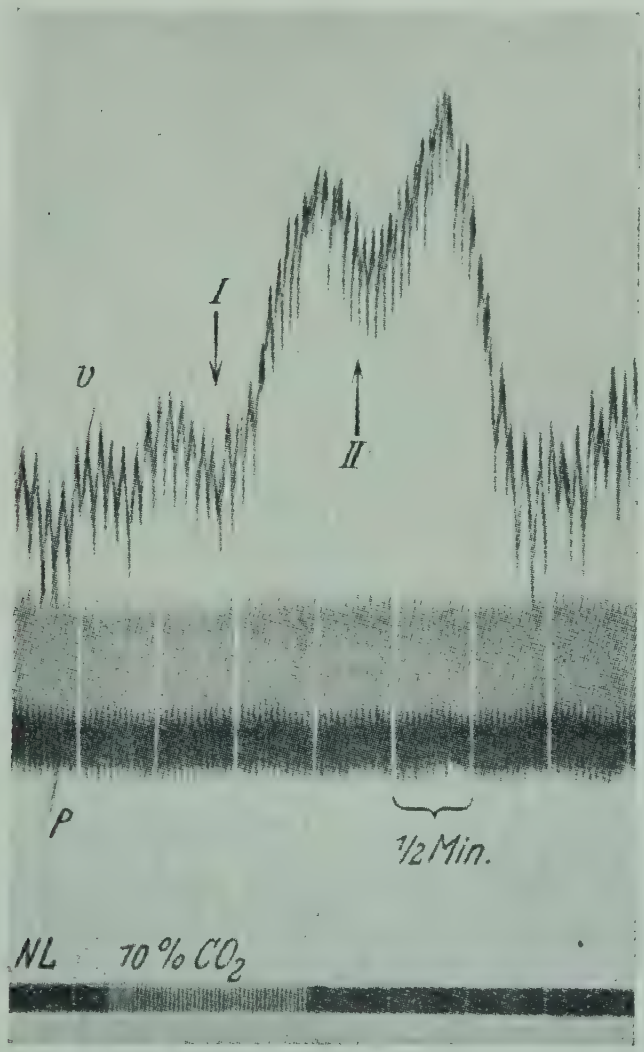


FIG. 425. Record of Velocity of Blood flowing along the Superior Mesenteric Artery of a Cat. The artery was perfused with blood from a pump, the perfusing pressure being kept constant, and the blood being arterialised in a separate pair of lungs. Upper tracing = blood velocity : the sudden downward strokes, about fourteen per minute at the commencement, are produced by a device which shows when each 1 c.c. of blood has passed through. They get closer together at the arrow II., and the whole tracing is moved upwards, both of which events show that the blood is flowing more rapidly. Second tracing : P = perfusion pressure, and time tracing in half minutes. Lower tracing shows the time during which the lungs of the perfusion circuit were supplied with a mixture containing 10 per cent. CO_2. (FLEISCH *et al.*, *Pflüger's Arch.*, 1932, **230**, 814.)

on the vessel is heated at a known rate (by a high-frequency current) and the rise of temperature is then recorded a short way downstream by a thermoelectric couple. The rise varies inversely as the velocity of flow. A method based on an electromagnetic principle has also been described.*

The velocity of the blood in the capillaries can be measured by direct observation of the capillaries under the microscope, and noting the time it takes for a blood corpuscle to move a measured distance.

These methods give mean velocities and give no information of the rapid changes occurring in the velocity of the blood at each heart beat. For this purpose we must have recourse to some such instrument as the *differential Stromuhr* used by Fleisch.† This is based on an application of Venturi tubes. If a current of blood, rendered incoagulable by a suitable method, be directed along the tube AB (or CB) possessing two vertical side tubes (Fig. 424), the pressure in the first branch will be greater than that in the second, since when the flow is from A to B, the momentum of the moving mass of blood is added to the lateral pressure of the fluid; if from C to B, there is still a difference, though a smaller one, due to the lateral pressure gradient along the tube. In recording the variations in blood flow, *e.g.* in an artery, the apparatus is inserted in the artery, so that blood enters it at A and leaves it at B to continue on its course along the vessel.

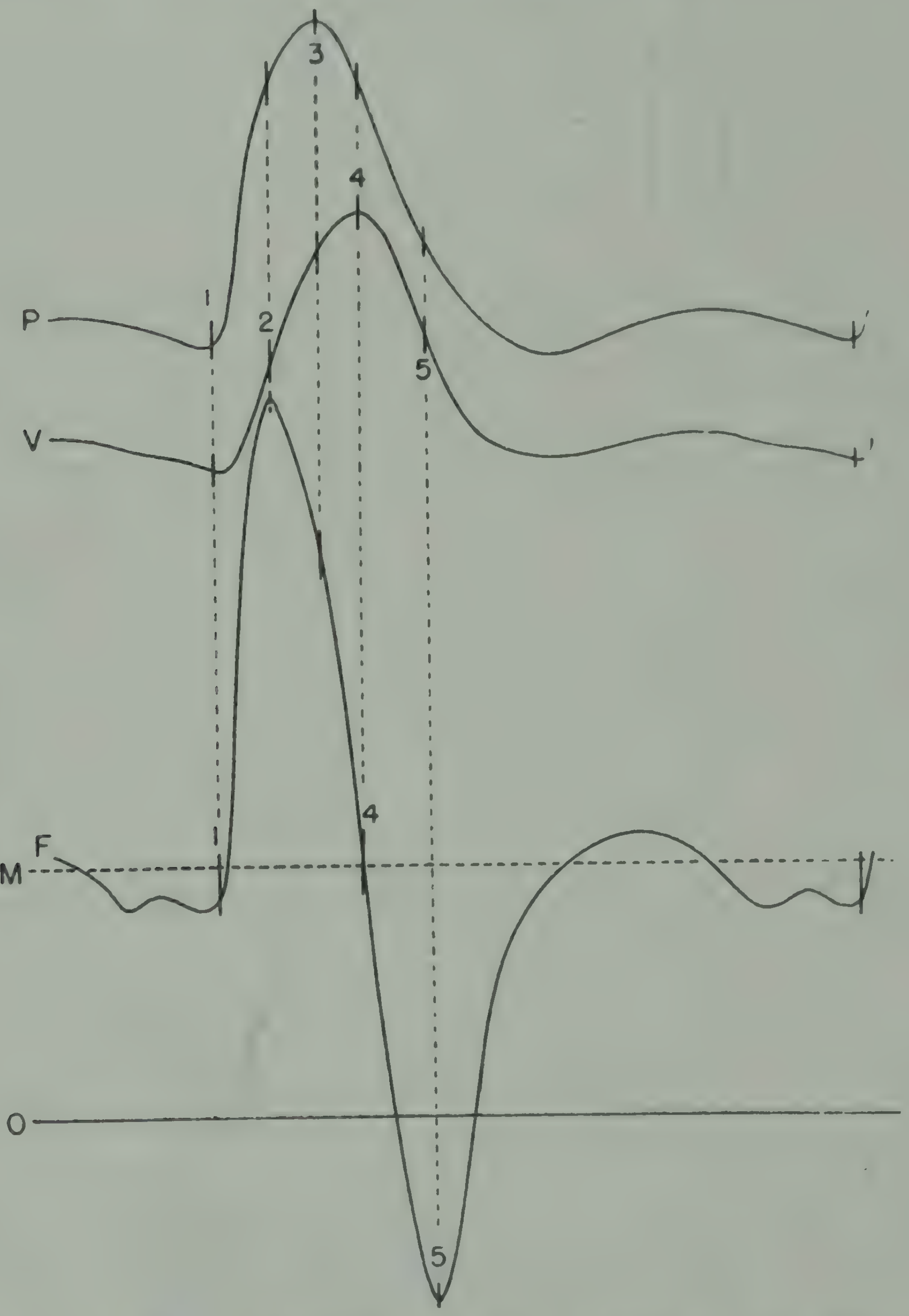


FIG. 426. Femoral artery pulsations in dog. Curve P, the pressure changes; curve V, the intravascular volume curve; curve F, the flow pattern. Line M, mean flow level; line O, zero flow level. Line 1, onset of flow and pressure rise: 2, maximal rate of forward flow and volume increase; 3, peak of applied pressure; 4, point of maximal intravascular volume; 5, maximal rates of back flow and volume decrease. (SHIPLEY, GREGG and SCHROEDER. *Amer. J. Physiol.*, 1943, **138**, 718.)

The difference of pressure will be greater if the flow accelerates, and will disappear if it ceases. It can be recorded by means of a membrane manometer, one surface of which is in connection with one branch and the other with the other branch, as shown in Fig. 424. The instrument must, of course, be calibrated empirically if the actual velocities are required. Fig. 425 gives a tracing taken with such an instrument, and illustrates the increase of blood flow which results when the splanchnic area is perfused, at constant pressure from a pump, with blood containing an increased amount of carbon dioxide. The velocity in the artery shows regular fluctuations with each stroke of the pump, as it would with the heart beat, an increase with each systole, and a decrease with each diastole (the tracing is on too slow a surface to show these clearly); there are also slower fluctuations in the form of waves (six are seen in the tracing) which are due to rhythmic changes in the vascular tone.

---

* KOLIN. *Amer. J. Physiol.*, 1937, **119**, 355.

† FLEISCH. *Pflüger's Arch.*, 1920, **178**, 31.

The Arterial Velocity Pulse. During systole the velocity of the blood in any part of the arterial system is greater than during diastole; thus in the carotid artery the velocity during systole may be as great as 500 mm./sec., and during diastole only at the most two-thirds as great. The fluctuations in velocity attending each heart beat are only approximately parallel to the fluctuations of pressure which we know as the pulse. By the use of accurate recording flow meters capable of following rapid changes in the velocity of the blood, it has been shown that in the large arteries, such as the axillary, femoral and carotid arteries, the peak velocity corresponding with the period of expulsion from the left ventricle is followed by a rapid decline in speed, culminating in a brief back-flow.* The smaller arteries, *e.g.* renal, hepatic and superior mesenteric, do not show this back flow, and in the smallest arteries, capillaries and veins of small size, there is little in the way of flow pulsations. The relations between pressure, velocity and volume, shown in Fig. 426 are complex, and cannot be discussed here: it will be seen, for instance, that the volume change does not follow the rapid pressure changes, and this is ascribed to the effect of viscous damping. What these valuable tracings do show, however, is that great care needs to be taken in the construction of apparatus for hæmo-dynamical research, and in the interpretation of the results obtained.

(3) **The Mean Volume Flow.** The volume of blood which flows through any portion of the circulatory system in a given time bears no simple relation to the actual speed of movement of the blood along the vessels concerned. For whereas the volume of blood transmitted at constant arterial pressure varies directly as the fourth power of the diameter of the small arteries that supply the part, and inversely as the resistance of the total circuit, the speed of movement at any given point under a constant head of pressure will vary directly as the volume transmitted, and inversely as the total sectional area of the local bed, at that point. It usually follows, therefore, that any dilatation of the arterioles supplying an organ leads to an increase in the velocity of movement of the blood, as well as to an increased input of blood to the organ.

The flow may be measured directly by actual collection of the blood, rendered incoagulable by heparin, as it emerges from a vein; the time required to collect a given volume is measured and the blood re-injected into the circulation. Or we may measure the blood entering by an artery, by inserting a stromuhr of some pattern on its course.

Another method is by the use of the plethysmograph (p. 643), by which the volume of an organ is recorded. If an organ is placed in a plethysmograph, and the outlet of blood by the veins be then obstructed for a definite *short* time, the organ will swell by an amount that can be read off from the record, so giving the amount of blood that entered it during the period of occlusion.† The flow of blood through the hand or arm in man can be arrived at by this method, which has been widely employed.

The method for the forearm is illustrated in Fig. 427. The forearm is placed in a plethysmograph P, connected with a suitable volume recorder R, and filled with water at an appropriate temperature. Above and below the plethysmograph are placed sphygmomanometer cuffs: the upper one V, for compressing the venous return, the lower one A, for shutting off the hand. In making a measurement, first A is inflated to a pressure, say 200 mm. Hg. well above the arterial; then V is suddenly inflated to a pressure, say 70 mm., which is above venous, but below arterial pressure. The forearm swells with inflowing blood. After 10 to 15 secs. both cuffs are deflated, and the recorded expansion measured against a calibration of the volume recorder. By a slight modification, the flow into the hand can be measured. Flow into forearm mainly represents muscle flow, that into the hand flow through skin.

---

* Shipley, Gregg and Schroeder. *Amer. J. Physiol.*, 1943, **138**, 718.

† Brodie and Russell. *J. Physiol.*, 1905, **32**,, Proc. xlvii; Lewis and Grant. *Heart*, 1925, **12**, 73.

With a cardiac output of 5·5 l./min. and body weight of 70 kg., the mean

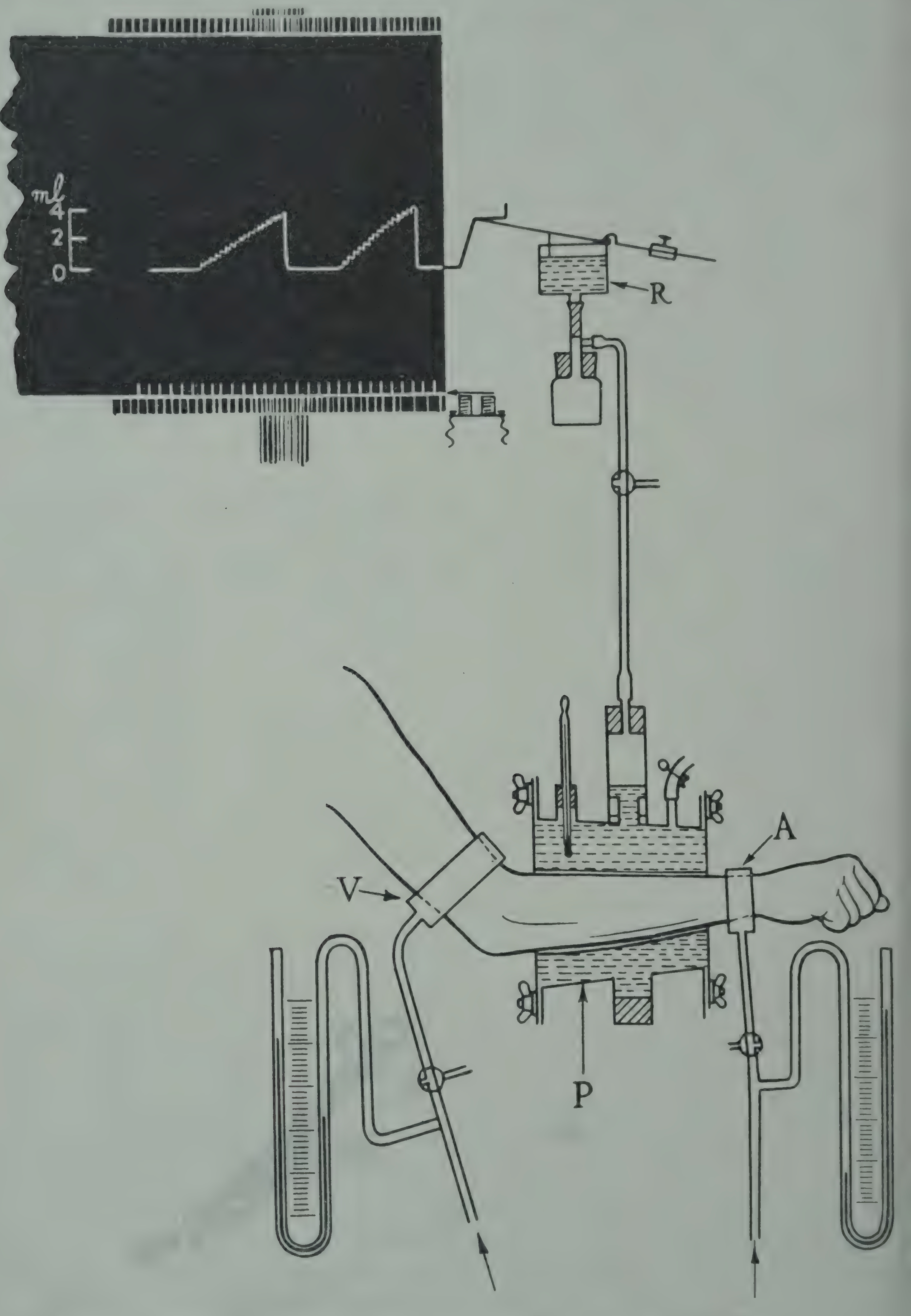


Fig. 427. Illustrating method of measuring forearm blood flow.

flow for the whole body would only be 8 c.c./100 g./min. The volume of bloo flowing through the various organs shows wide variation, depending chiefly o the intensity of metabolism of the organ, and so on its oxygen requirement If the organ becomes more active the blood supply to it is increased by dilatatio of the arterioles, or capillaries, or both.

The following figures give approximate rates of flow per 100 g. organ per minute :

Thyroid	560 c.c.
Kidney	400 c.c.
Liver (1/3 arterial, 2/3 portal)	150 c.c.
Heart (coronary circulation)	100 c.c.
Brain	55 c.c.
Intestine	70 c.c.
Resting muscle	3 c.c.
Tissues of hand (mainly skin)	4 to 30 c.c.

Of the total output of the left side of the heart it is probable that in states of muscular rest about half goes to supply the head, upper extremities and liver [*] : about one-quarter goes to the kidneys, and a quarter to musculature and other regions.

THE ARTERIAL PULSE

The systolic rise of pressure at each heart beat causes an expansion, spoken of as the pulse, which can be felt in any superficially situated artery, such as the radial. As we have seen, the amplitude of the pulse decreases as we go farther away from the heart.

From time immemorial the physician has sought, by feeling the pulse, to come to some idea as to the condition of the heart. A number of different qualities have therefore been distinguished, *e.g.* the *rate* may be rapid or slow, the size or *amplitude* large or small, while the *hardness* of the pulse is determined chiefly by the blood pressure.

If the pulse is compressible it is spoken of as *soft ;* if it can only be obliterated with difficulty it is *hard.* Certain combinations of these qualities are also described. Thus a large and hard pulse is spoken of as *strong*, a *weak* pulse being both small and soft. If the rhythm of the heart beat is irregular the pulse is also *irregular.* An *intermittent* pulse is one in which one heart beat is dropped occasionally, *e.g.* once in every four or eight beats, and may be due either to the interposition of a ventricular contraction which is too weak to propel the pulse so far as the radial artery, or to the occurrence of a premature contraction, followed by a compensatory pause.

A distinctive pulse is that known as the " *water-hammer* " *pulse,* which is observed in cases where the aortic valves are diseased and allow of regurgitation. The systolic rise of pressure in the arterial system is followed by an extremely rapid fall. Under such conditions, if the arm be held above the head, the pulse in the arteries of the wrist is felt as a smart blow.

Cause of the Pulse Wave. If the arterial system were perfectly rigid, the increased pressure at each ventricular systole would occur practically simultaneously at every point. The arteries are, however, elastic and distensible, so that the first effect of the flow of blood into the aorta is to distend the section of the aorta nearest to the heart. The shock due to this suddenly increased tension is transmitted from segment to segment of the arteries in the form of a wave, at the velocity of about seven metres per second.

It is important not to confuse the velocity of the pulse wave with that of the blood flow ; the latter is never greater than 0·5 metre per second, and is very much less than this in the smaller arteries. Perhaps the difference between the two quantities may be made clearer by analogy. When a railway engine shunts a row of slowly moving trucks, a jolt is very quickly passed on from truck to truck until the end van is reached, though in that short time the train as a whole has moved a very short fraction of its length, by the constant push of the engine. In the arteries, the continuous pressure is furnished by the elastic reaction of the arterial wall, and we see how the impact of the blood may travel quickly as a wave of increased pressure, while the blood itself is moving slowly along, impelled by the reaction of the arterial wall.

---

[*] Levy and Blalock. *Amer. J. Physiol.*, 1937, **118**, 368.

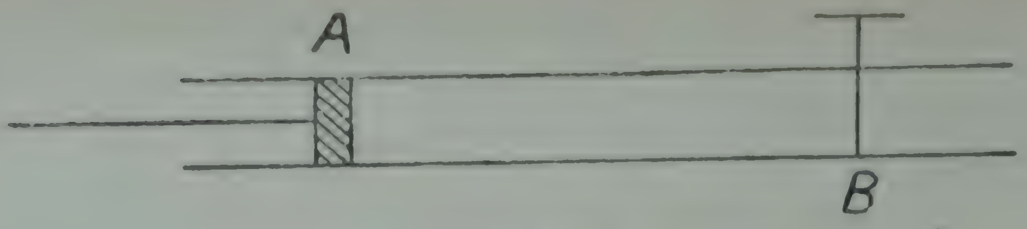


FIG. 428

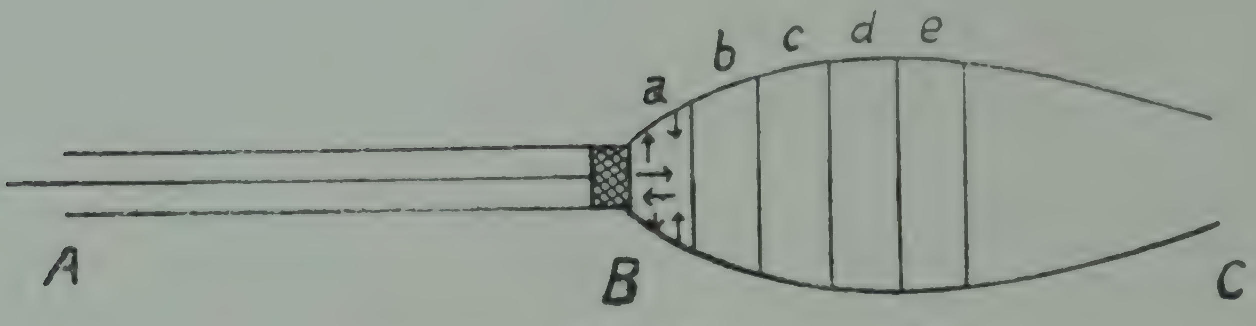


FIG. 429

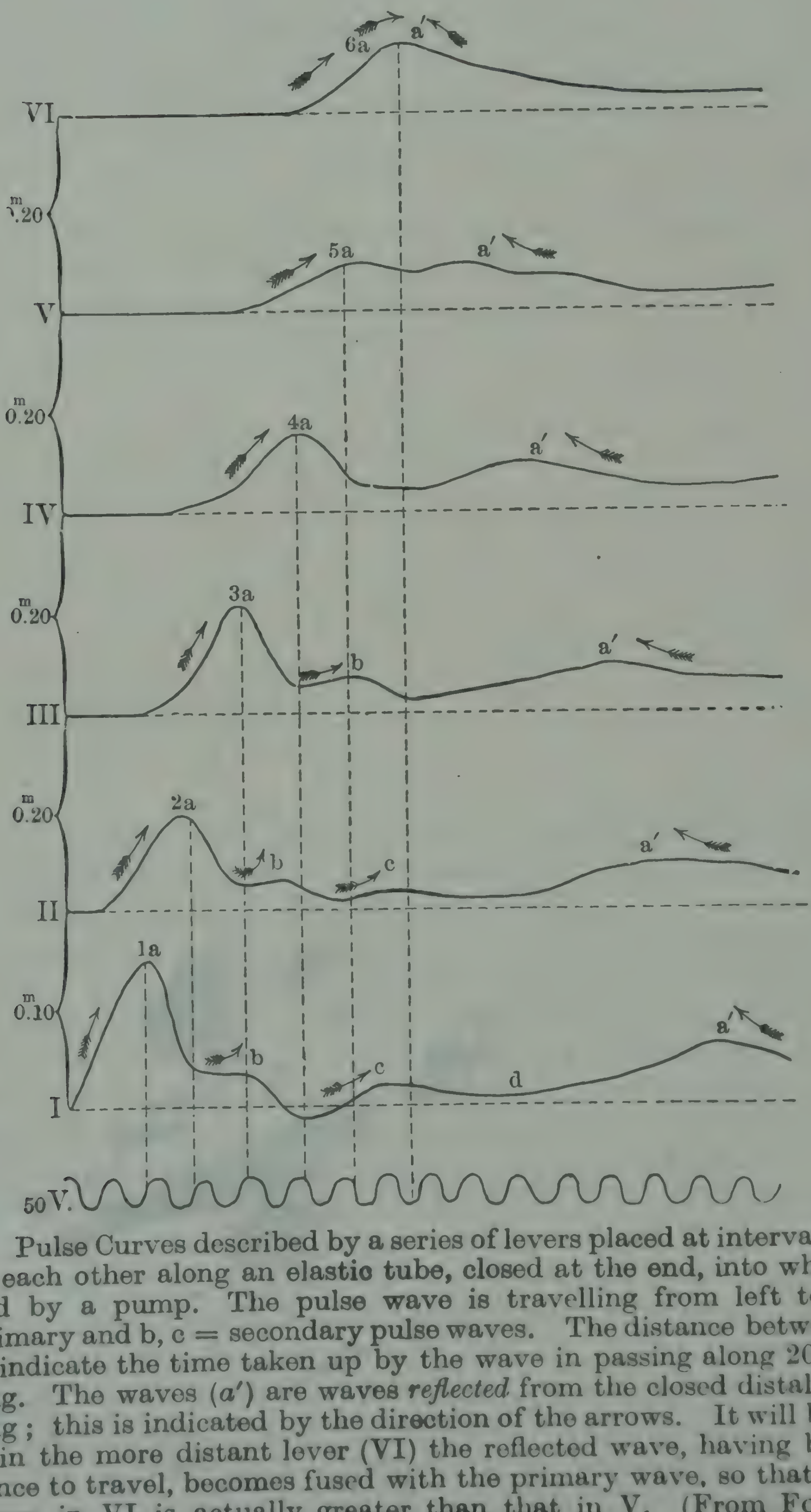


FIG. 430. Pulse Curves described by a series of levers placed at intervals of 20 cm. from each other along an elastic tube, closed at the end, into which fluid is forced by a pump. The pulse wave is travelling from left to right; a = primary and b, c = secondary pulse waves. The distance between vertical lines indicate the time taken up by the wave in passing along 20 cm. of the tubing. The waves (a') are waves *reflected* from the closed distal end of the tubing; this is indicated by the direction of the arrows. It will be observed that in the more distant lever (VI) the reflected wave, having but a slight distance to travel, becomes fused with the primary wave, so that the rise of pressure in VI is actually greater than that in V. (From FOSTER, after MAREY.)

In a rigid tube AB (Fig. 428) provided with a piston at the end A, and filled with fluid, an inward movement of the piston at A will cause a simultaneous outflow of fluid at the end B. If the end B is closed, the piston at A cannot be moved at all. Pressure applied to the piston will raise the pressure simultaneously at all points in the tube AB. The increased pressure applied at A is therefore transmitted with practically no loss of time to all parts of the tube AB. The transmission will be retarded if we have a tube whose wall is distensible and elastic. If we imagine (Fig. 429) an elastic tube BC filled completely with water and connected at B to a rigid tube, which is provided with a piston, the first effect of a rapid movement of fluid driven in by the piston will be a rise of pressure at the point immediately in front of the piston, viz. at *a*. The wall being distensible, and pressure being propagated along the fluid in every direction, the rise of pressure at *a* will be spent partly on the particles of fluid in front of it, viz. at *b*, but also on the walls of the tube, so that this is stretched and the cross section of the tube enlarged. The distended segment at *a* will then exert a pressure on the contained fluid, driving this backwards and forwards. The fluid on its side towards the piston will tend to come to a stop, while that towards the distal end of the tube will be accelerated. The distended wall therefore returns to its original diameter, and the next segment at *b* is stretched in its turn, so that a wave of increased pressure is propagated along the tube in the direction of the arrow.

The velocity with which this wave is propagated depends on the density of the fluid, *i.e.* its inertia, and on the resistance of the walls of the tube to distension, *i.e.* on the rapidity of its recovery. The velocity of propagation of the wave of increased pressure, or the wave of expansion of the artery, is expressed by the following formula : *

$$v = \frac{3 \cdot 57}{\sqrt{E}}$$

where v = velocity in m./sec., and E = per cent. increase in volume of vessel per mm. Hg increase of pressure.

The elastic coefficient of the arterial wall becomes larger as the arterial wall is distended. The velocity of propagation of a wave of pressure in the arterial system will thus be greater when the arterial pressure is high than when it is low.

If the end C of the tube is closed, the wave of positive pressure on arriving at C will be reflected back as a positive reflected wave. If a tracing be taken of the oscillations or variations of pressure in the tube, two waves at least are seen, one of which is the primary wave due to the movement of fluid caused by the piston ; the other is the secondary wave reflected back from the periphery. That the secondary wave is a reflected one is shown by the fact that the nearer to the closed end the pulse is recorded, the nearer is the secondary to the primary wave, as is seen in Fig. 430.

If the tube BC be widely opened at C, a reflected wave is also observed, but this time the wave is one of reduced pressure. The production of this wave is dependent on the momentum of the moving column of fluid. If in the tube *ab*, with a tap at *c* and a manometer *m* (Fig. 431), the current of fluid be suddenly checked by turning the tap *c*, the column in front of the tap, having a certain momentum, will tend to go on moving and therefore produce a suction or "negative pressure" behind it. When a wave of pressure arrives at the open end of a tube, there is a sudden drop in the resistance to the output, and the momentum of the mass of fluid which is thrown out causes a similar suction or "negative pressure," which

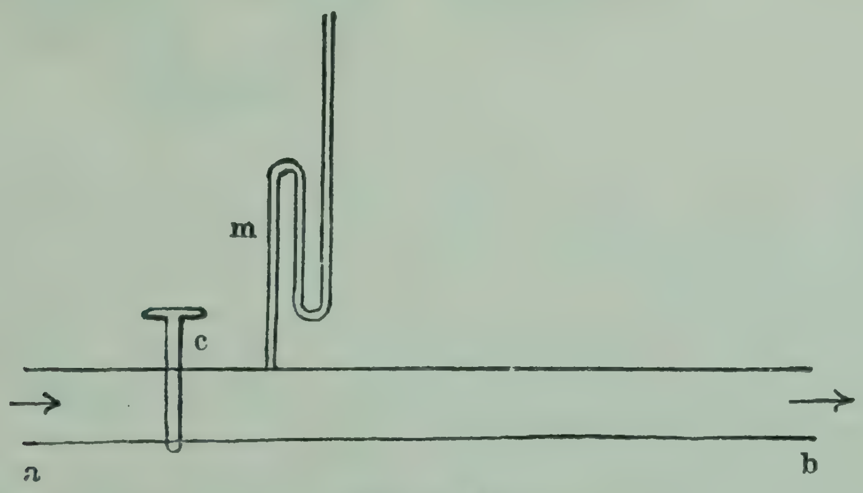


FIG. 431.

---

* BRAMWELL and HILL. *Proc. R. S.*, 1922, **93B**, 298.

travels back the whole length of the tube. If the end of the tube is only partially closed, every primary positive wave will be transformed into a reflected one which is partly positive and partly negative. Since both these reflected waves travel through the tube with the same velocity and will mutually interfere, the result may be either a positive or a negative wave or nothing at all, according to the degree of constriction.

In a branching system of tubes, such as the arterial system, reflection of waves must take place at every division. All the conditions for the origination and interference of such waves are present in the arterial system. It is impossible, however, to say *a priori* whether any reflected wave will form a marked feature on the pulse tracing. It is possible that the multitudinous reflections which must occur in every part of the arterial system may interfere with one another to such an extent that they mutually annul each other. The identity of any secondary wave in the pulse tracing must therefore be determined by exact measurement and experiment.

Pulse Tracings. To study the pulse it is necessary to obtain a graphic record of the expansion of the arteries or of the changes in pressure which produce this expansion. The former is the more accurate because, owing t the fact that the extensibility of an artery diminishes with rising pressure, given increment of pressure causes a greater expansion at the commencemen of the rise than towards its peak. For accurate records, a manometer whic has very little inertia, such as Wiggers' (p. 557), must be used, or the expansio of the artery may be recorded photographically. The expansion of the arter can be approximately registered by means of a lever, which may be made t rest upon an artery, *e.g.* the radial, and the movements of which are record on an appropriate surface. Such an instrument is called a *sphygmograph*.

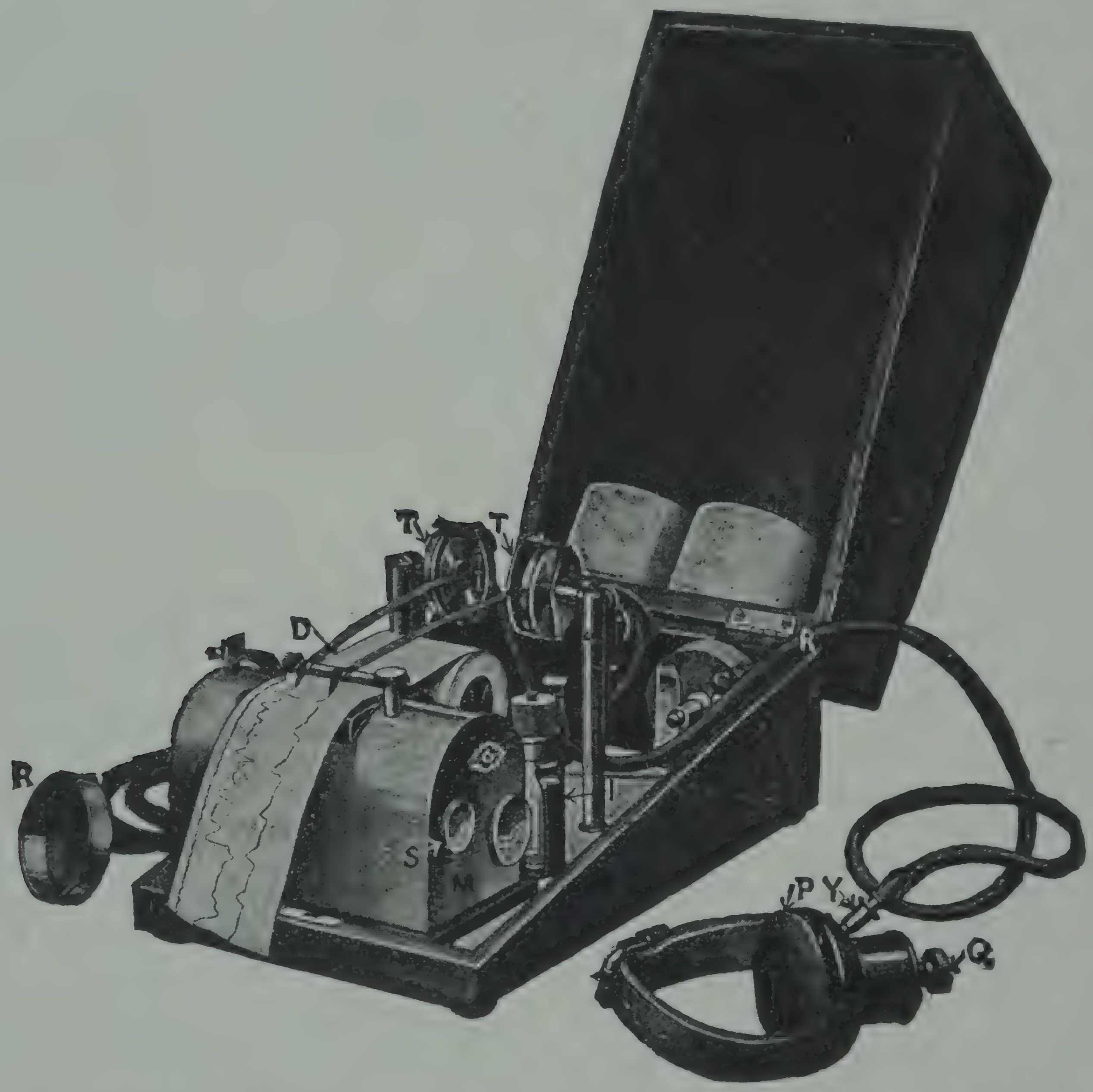


FIG. 432. The Polygraph. The record is made on the paper, wound out by clockwork from a roll, by the ink pens D. These are attached to levers of the tambours T-T, which are actuated by air-transmission from the receivers. P is the glycerin pelotte which can be strapped round the wrist; the pulsations of the radial artery cause movements of the air along the tube leading to the tambours, and so records the pulse tracing of the artery. R, an open capsule, can be used for application over the jugular vein, in recording the venous pulse. (CAMBRIDGE INSTRUMENT CO. LTD.)

Alternatively, one of the various forms of *polygraph*, depending on air transmission by tambours, may be used. They have the advantage that records from two or three sources, *e.g.* from two different arteries, or from artery and vein, or cardiac impulse, etc., may be made simultaneously (Fig. 432).

In all such instruments the moving parts have considerable inertia, so that the curve they give is always deformed. This fact must be borne in mind when comparing the pulse curves obtained by means of a sphygmograph or polygraph with those given by the more perfect forms with optical methods of recording. If an open cup be firmly pressed to the skin over an artery, such as the carotid or subclavian, and be connected by a tube with a "segment capsule," the excursions of which are recorded photographically, curves may be obtained reproducing in all particulars those given by optical manometers (Fig. 433).

The pulse wave undergoes considerable modification in shape and time relations as it passes along the arterial system. The time-pressure or time-volume curves taken in the aortic arch or its great branches, such as the subclavian or carotid, closely follow the time-pressure curve of left intraventricular pressure during the phase of ejection, and so are of importance in man as an index of the action of that chamber. This type of pulse is called the *central arterial pulse* to distinguish it from the *peripheral arterial pulse* as recorded in smaller arteries, such as the radial.

The differences between a *central* pulse as recorded in arteries such as the subclavian, and an intermediate or terminal pulse given by arteries such as the carotid, radial or dorsalis pedis * are due to the fact that the contour of the pulse, as it occurs at the aortic arch, is modified in its transmission to the periphery by friction, by incidental vibrations and by interference due to reflected waves. The central pulse, as recorded in Fig. 433 from the subclavian artery, shows the following features :

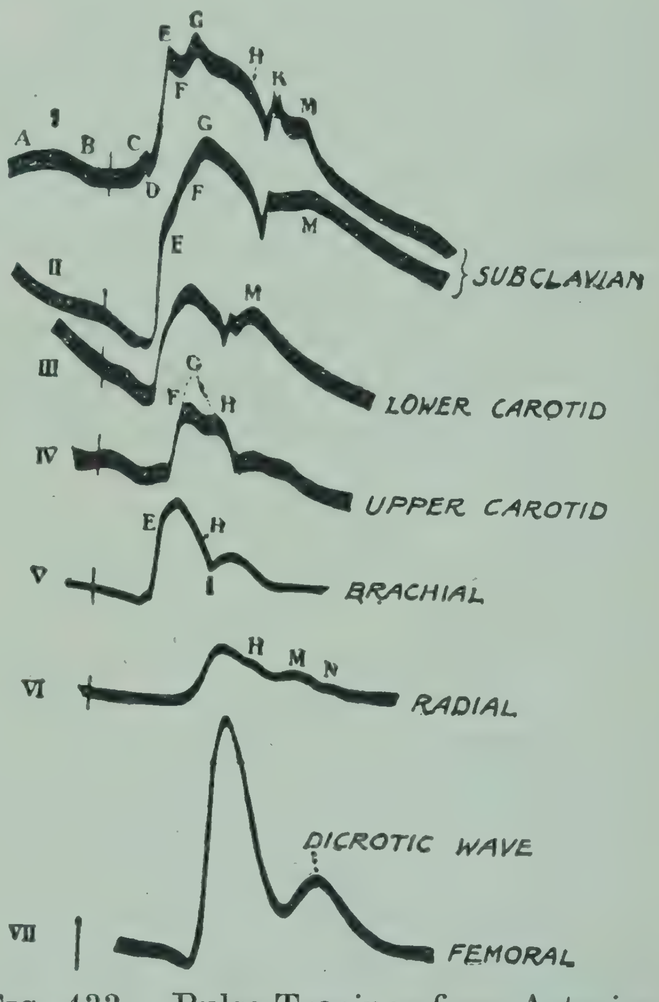


Fig. 433. Pulse Tracings from Arteries, taken by Air Capsules and Optical Recording. (Wiggers.)

(1) An oscillation between A and B, due to the auricular systole (transmitted).

(2) A second elevation between B and D which corresponds with the isometric part of the ventricular contraction. (1) and (2) are probably vibrations transmitted across the thorax.

(3) At D the semilunar valves open, and blood is rapidly ejected into the aorta, giving the main pulse between D and G. The sudden rise of pressure in the aorta sets the column of blood within the elastic vessel into vibration, giving the overshoot at E and the return wave or anacrotic halt, E F.

(4) The arterial curve now follows the intraventricular pressure (*vide* Fig. 364, p. 559), the flow of blood continuing quickly up to G, but then slowing off during the rest of the ventricular systole to H (Fig. 433).

(5) At the beginning of ventricular diastole there is a rapid backward movement of the blood towards the heart, causing the *incisura*, H I, and brought up rapidly by the closure of the aortic valves. The rebound of the blood from the valves causes one or two secondary elevations, K, after the incisura.

(6) The curve then falls gradually until the next ventricular systole supervenes, sometimes with slower waves at M.

The peripheral pulse gives similar features to the central pulse, but modified by the conditions mentioned above. The preliminary and anacrotic oscillations disappear (though the anacrotic oscillation may be counterfeited by a momentum swing of the lever in the ordinary imperfect sphygmograph); the incisura becomes less sharp, and is replaced by the *dicrotic* dip, and the wave K is replaced by the sometimes very conspicuous *dicrotic* wave; the main upstroke is delayed and its rise more gradual, while the amplitude of the whole curve is usually diminished. These changes are more marked as the periphery is approached, until finally in the smallest arteries and in the capillaries the pulse may disappear altogether.

---

* Wiggers. *J. Amer. Med. Ass.*, 1915, **64**, 1485.

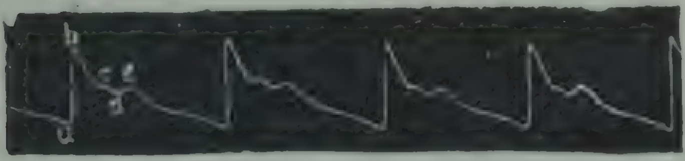


FIG. 434. Pulse Curve from Radial Artery.

Fig. 434 represents a pulse curve taken from the radial artery by means of a sphygmograph. It shows the anacrotic limb, or percussion wave, *a–b*, a peak at *b*, and a slower downstroke, or catacrotic limb which is broken by secondary waves, the most conspicuous of these being the *dicrotic wave*, with its preceding *dicrotic notch*, *c*. This sphygmogram should be compared with the curves, taken by more perfect methods, in Fig. 433, so as to appreciate the deformation by ordinary sphygmographs of the real curve of expansion of the arterial wall.

The pulse curve gives approximate information as to the condition of both the heart and the arterial system. Thus, the period between the beginning of the primary elevation and the dicrotic notch corresponds to the outflow of blood from ventricle to aorta. A large pulse curve does not necessarily indicate a big output, since the expansion of the artery is determined not only by events occurring in the aorta but also by the expansibility of the artery and the resistance in the peripheral branches.

The general form of the pulse curve varies with changes in the heart, in the arteries and in the peripheral resistance. Thus, some sphygmograph curves may present secondary elevations on the ascending part, and are called *anacrotic*.

A curve of anacrotic type is normally seen in the subclavian artery (Fig. 433), but in the radial artery usually only in cases in which, owing to aortic stenosis, there is an ascending plateau in the intraventricular pressure tracing. When the peripheral resistance is high, or when an extra large amount of blood is thrown into the aorta at each stroke of the heart (*e.g.* by slowing of the rate), the aortic pressure rises very abruptly at the commencement of the ejection phase, and then more slowly so long as blood is flowing in, and we get an ascending systolic plateau and an anacrotic pulse. If the pulse rate is quickened the anacrotic wave disappears.

The production of the dicrotic elevation is favoured by any influence which increases the distensibility of the arteries or causes the primary elevation of the pulse to be extensive and sharp. Thus it is much more pronounced in young people, and in aortic incompetence ("water-hammer pulse"). When the peripheral resistance is low through relaxation of the arterioles, and the heart is beating forcibly, as in fever, and also to some extent after alcohol, the dicrotic elevation becomes so large that it may be felt at the wrist, and the mistake has been committed of taking the dicrotic wave for a normal beat, and so doubling the actual rate of the pulse.

Reflected Waves. It has been suggested that, in the production of such a marked dicrotism, reflection from the periphery may play an important part. With a high blood pressure and rigid arteries, a reflected wave will travel back very quickly and will tend to add itself to the primary wave. With a low blood pressure, dilated arteries, and the output of the heart thrown rapidly into a relatively empty arterial system, the primary wave will rise and fall very rapidly, and the reflected wave will travel back along the arteries more slowly, so that its main effect might be to add to the dicrotic elevation normally proceeding outwards from the heart towards the periphery. The Fig. 430 VI would represent the condition as it is found in the femoral artery under normal circumstances when, according to Frank, the reflected wave adds to the height of the primary wave. In Fig. 430 V the reflected wave "a" would tend to add to any dicrotic elevation present at this point, and might represent the relation existing in the arterial system with relaxed arteries and a heart beating forcibly, but throwing out only a small amount of blood at each beat.

It would seem that the central pulse can be referred entirely to events taking place in the heart during systole, or at the beginning of the aorta at the commencement of diastole; and there is no reason to assume the co-operation of waves reflected from the periphery to explain any of the secondary waves observed. The effect of the dispatch of the sudden wave of pressure started in the aorta along an elaborate system of elastic tubes must be to diminish the rapidity of onset of each primary wave, and therefore to diminish the secondary vibrations of the curve. In an elastic system of tubes such as the arterial system, there are factors at work analogous in many respects to those responsible for the deformation of the curve given by an imperfect manometer.* These would be of two kinds—viz. oscillations of the column of fluid within the stretched arterial wall, and the reflection of waves from different points in the periphery. Many of these reflections will interfere with and annul one another. But in the arterial system there are certain points from which distinct reflections of waves can be expected—*e.g.* in the circle of Willis, at the bifurcation of the aorta into the two iliac arteries, and in the superficial and deep arterial arches in the hand and the foot. We have distinct evidence that such waves are set up and modify the form of the pulse in the femoral and brachial arteries and their branches. Thus in Fig. 433 V-VII, the primary rise of pressure in the femoral artery is higher than even the primary rise in the subclavian. This condition of things is explicable only on the assumption of a reflected wave passing back along the artery just after the passage of the primary wave, so that the two are summated. In the same way, although the dicrotic depression in the curve (Fig. 433, VII) is no doubt mainly the propagated effect of the incisura observed in the aortic pulse, it is probably deformed, and the subsequent elevation exaggerated, as a result of reflection of some wave from the periphery. The occurrence of reflected waves may serve to explain why the systolic pressure in the femoral artery is found higher, and the diastolic pressure lower, than in the brachial. The femoral artery being more rigid than the brachial, and the peripheral resistance more definitely localised, reflected waves occur in the artery at so short a time after the primary wave has passed down, that there is summation of the two waves, with production of a higher maximum and a lower minimum than was present in the waves as started in the aorta. If the leg be plunged into hot water, so as to dilate all its arterioles, this difference between the arm and the leg systolic pressures disappears. The varying development of reflected waves on the two sides may also explain why the systolic pressures in the two arms are rarely found to be identical.

It has been pointed out † that, as the pressure rises rapidly in the artery at each pulse and the wall becomes more rigid, there is a corresponding increase in the velocity of propagation of the wave along the artery which would tend to make the rise steeper towards the periphery. This will be more marked the more rigid the arterial wall, and may account for the primary elevation in the big arteries of the legs being higher than in the more extensible arteries nearer the heart.

As the pulse wave travels along it changes in form, and under certain conditions part of the wave front becomes steeper, and eventually unstable. The result is a turbulent motion which may be compared with the formation of breakers in water waves on a shore. Such a change gives the peculiar character to the " water-hammer pulse " and is also responsible for the cracking character of the Korotkov sounds when the pressure on the armlet is just above the diastolic pressure.‡

The *velocity of the pulse* wave can be measured by taking simultaneous tracings from two arteries separated by some distance from one another, such as the femoral artery and the dorsalis pedis, or from the carotid and radial arteries. In healthy individuals the velocity varies between 5 and 9 metres per second. The more rigid the arteries and the higher the arterial pressure the greater will be the velocity, so that the velocity of propagation gradually increases with advancing age,§ and is higher in the arteries of the lower extremities than in the more distensible arteries of the arm.

The length of the pulse wave can be found by multiplying the velocity of transmission by the time occupied by the wave in passing any given point (0·8 sec.), so that if the velocity of transmission be taken as 7 metres per second, the length of the wave is about 5·6 metres. The pulse wave thus begins at the periphery long before it has been completed in the aorta.

---

* FRANK. *Z. f. Biol.*, 1905, **46**, 441.

† BRAMWELL and HILL. *J. Physiol.*, 1923, **57**, Pr. lxxiii.

‡ BRAMWELL. *Lancet*, 1937, ii, 366.

§ BRAMWELL, HILL and MCSWINEY. *Heart*, 1920, **10**, 3; BRAMWELL and HILL. *Proc. Roy. Soc.*, 1922, **93B**, 298; HICKSON and MCSWINEY, *J. Physiol.*, 1924, **59**, 217.

The Volume Pulse in Organs. Each heart beat, by causing a pulsatile expansion of the arteries throughout the body, results in the organs momentarily receiving blood faster than the veins carry it away. Hence there is a pulsating distension of all the vascular organs of the body, which may be graphically recorded by a delicate plethysmograph, one pattern of which, suitable for use with a finger, is shown in Fig. 435. The form of the volume pulse of the finger is seen to be similar to that of the radial pulse, though nearly 0·1 second behind it in time (Fig. 436).

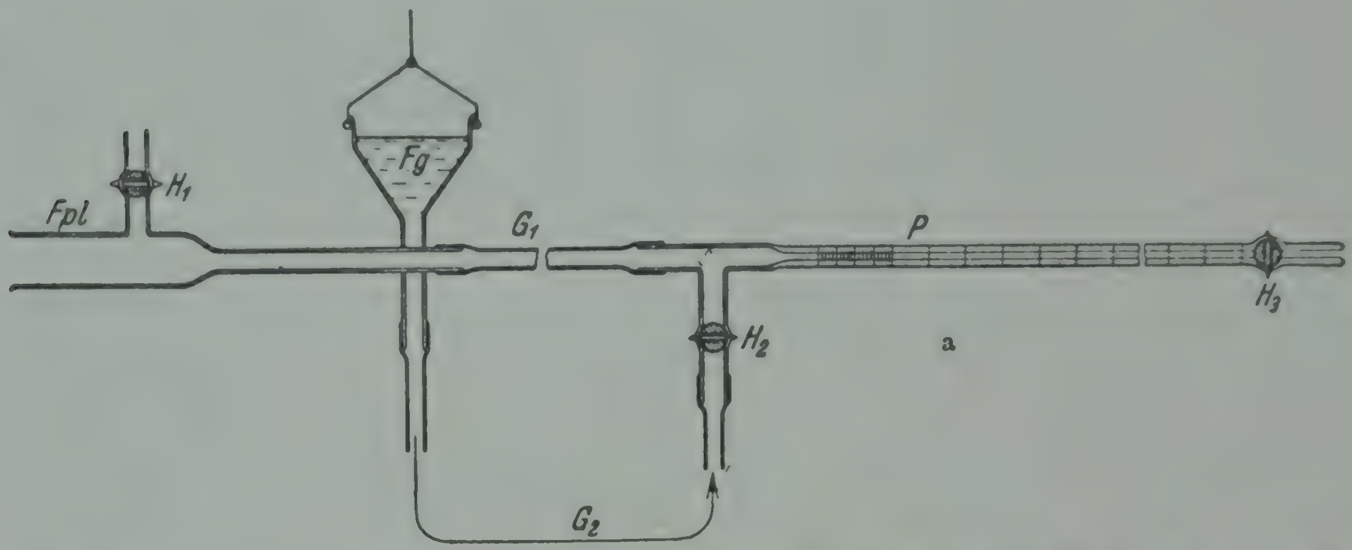


Fig. 435. Finger plethysmograph. The finger is placed airtight in *Fpl,* which is filled with warm water from *Fg.* The pulsations of the water meniscus in the capillary *P* are recorded photographically. (Goetz. *Pflüger's Arch.*, 1935, **235,** 271.)

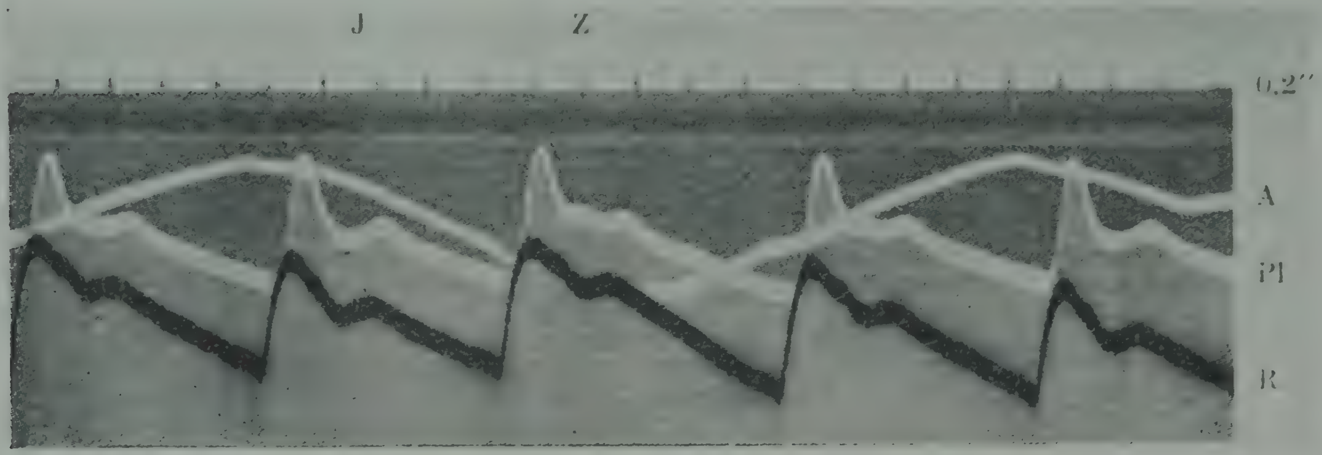


Fig. 436. Volume pulse of finger (Pl) recorded together with the radial pulse (R) and respiration (A) (inspiration upwards). Note the different character of the dicrotic waves at inspiration and expiration. (Goetz. *Pflüger's Arch.*, 1935, **236,** 271.)

When the tissues are tense with inflammatory œdema, the pulsatile distension causes the familiar throbbing pain because the tension is raised at each pulse wave, and pain fibres thereby stimulated.

For Reference

Bramwell (1937). The Arterial Pulse in Health and Disease. *Lancet*, 1937, ii, 239. 301, 366.

Hess. *Die Gesetze der Hydrostatik u. Hydrodynamik.* (*H'dbh. d. norm. u. path. Physiol.*, VII., 2, ii c/1, 2.)

Wiggers (1928). "The Pressure Pulses in the Cardiovascular System." (London: Longmans, Green.)

THE CIRCULATION THROUGH THE CAPILLARIES

The circulation through the minute surface vessels of the skin in man, and in other tissues can readily be observed when a microscopic examination is made by reflected light. The skin is for this purpose covered with cedar oil or liquid paraffin (Fig. 437).

The capillary circulation is best demonstrated by examining under the microscope the tongue of the frog or the web of the frog's foot, when a network of vessels is seen, consisting of small arteries, capillaries and veins. The flow of blood in the arteries is rapid, in the veins it is slower, and is slowest in the capillaries. In the capillaries the flow is from arteries to veins, though, on account of the reticular arrangement of these vessels, the direction of the stream through some of them is not quite constant. If a group of capillaries be watched for some time, the blood may at first hurry through a number of them with rapidity ; the flow then becomes slower and later may quicken up to a moderate pace again. These variations in the capillary flow are probably associated with spontaneous alterations in the condition of contraction of the small arteries supplying the group of capillaries. It is easy to observe that the arterial flow is pulsatile, the pulsation being reduced, or disappearing, in the capillaries and veins. Another difference between the circulation in these three kinds of vessels is to be found in the condition of the peripheral zone. In the arterioles the blood stream is divided into two parts, the peripheral stream, consisting only of colourless plasma with occasionally a stray leucocyte—and an axial stream, in which all the red blood corpuscles are being hurried along. In the veins there is a similar peripheral plasmatic zone, but here we find regularly scattered leucocytes which travel rather more slowly than the axial stream of red corpuscles. The formation of this axial zone is purely mechanical, and may be imitated in any fluid containing in suspension particles whose specific gravity is somewhat higher than that of the fluid. In the capillaries there is no separation of the two zones, since the lumen of these vessels is so narrow that often the corpuscles, passing in single file, are squeezed out of shape as they progress. The corpuscles are highly elastic structures, and may be seen to bend if they impinge on the dividing point of two capillaries, before they are finally swept off by the stream into one or the other of them.

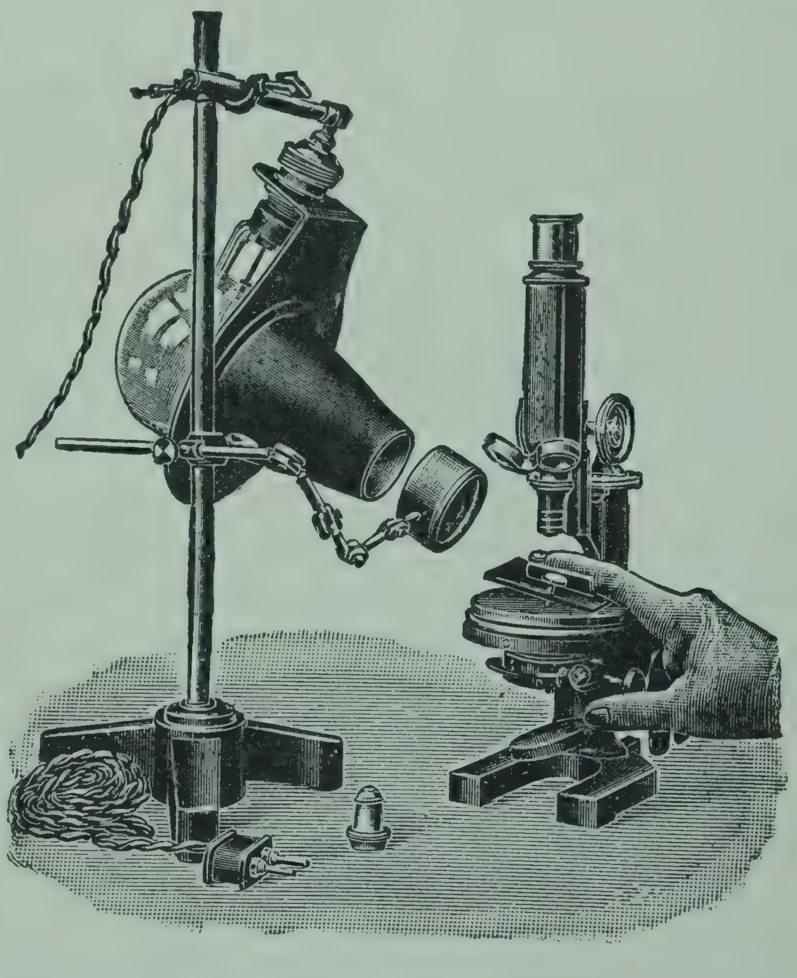


FIG. 437. Arrangement of Illumination for Viewing Skin Capillaries.

The capillary wall is composed of a single layer of flattened endothelial cells, which present little resistance to the passage through them by diffusion of dissolved substances. In this way the tissue cells obtain oxygen from the red blood corpuscles and nutriment from the plasma, and give off to the circulating blood carbon dioxide and other products of their metabolism. There is evidence that a sort of "contractility," due to a swelling of the nuclei of their endothelial cells when stimulated, is a universal attribute of the capillaries, and must play an important part in determining the amount of blood flow through an organ in accordance with its needs. In resting tissues many of the capillaries are so constricted as to allow of no flow at all.

The average length of a capillary is between 0·4 and 0·7 mm. The velocity of blood flow can be directly determined by observing under the microscope the time taken by any given corpuscle to travel a measured distance. The mean velocity determined in this way varies from about 0·5 to 0·8 mm. per second.

The blood pressure in the capillaries may be measured approximately by applying pressure to the outer surface of the skin or mucous membrane by means of a Recklinghausen capsule (Fig. 414) and noticing the point at which blanching of the surface commences. The errors of this method are considerable. The pressure in the capillaries, as found by this method, is low, and necessarily varies with the position of the part under investigation, *i.e.* with the hydrostatic pressure. It is raised, often greatly, when there is a dilatation of the arterioles, and lowered when these constrict.

Capillary pressures have also been measured in man by Carrier and Rehberg as follows: A fine glass capillary is filled at the point with saline solution and pushed into a skin "capillary" which is under observation under the microscope. Attached to the other end of the capillary tube is a water manometer, by means of which a definite pressure is applied to the column of saline solution in the capillary glass tube. When the blood capillary is pierced, blood will flow into the glass capillary if the capillary pressure is higher than that in the water manometer; if lower, no blood will enter: if the same, blood will pulsate in and out of the glass capillary at each heart-beat.

In man it has been found that the capillary pressure in the skin of the finger varies with the position of the hand with reference to the heart, and may be 40 cm. H_2O, or more, or as little as 4 or 5 cm.; but, however high up the hand is placed, it never falls below a certain level of about 4 to 5 cm. The reason for this is that the veins collapse and restrict the flow when the hand is held in such a position that the venous pressure becomes negative. The capillary pressure is thus always maintained at a positive value, which varies according to the position. In the feet the capillary pressure is normally prevented from attaining very high values, owing to the presence of valves in even the smallest veins, and to the occurrence of constant small muscular movements, which, by forcing the blood along these veins, enables them to act as a kind of pump. The pressure gradient along a systemic capillary is considerable. Landis* found the average pressure at the arterial end of skin capillary loops to be 32 mm. Hg and that at the venous end 12 mm. Hg. Pulmonary capillaries are wide, and their gradient small.

Owing to the fact that the resistance of the arterioles lies between the capillaries and the arteries, the pressure in the capillaries must stand in much closer relationship to that in the veins than to that in the arteries. One cannot therefore argue that a fall of arterial pressure necessarily involves a fall of capillary pressure in all parts of the body. We can only judge of changes in the capillary pressure by simultaneously taking the pressures in both the afferent and efferent vessels. If these both rise or fall together we may be certain that the capillary pressure also rises or falls. When the arterial and venous pressures move in opposite directions, it is difficult to say what alterations, if any, will be produced in the capillary pressure.

The resistance to the flow of blood through the capillaries is determined by the internal friction, *i.e.* the viscosity of the blood; as already explained (p. 618), when circulating in the blood capillaries, this is only about twice that of water. The velocity of flow through a capillary bed varies directly as the driving (arterial) pressure, provided this is above 20 mm. Hg, below which little flow occurs. The chief seat of the resistance in the vascular system is in the arterioles, and it is in this region that the chief fall of pressure occurs.

No part of the circulation shows greater variations than the capillary system. We must think of this as a vast irrigation system—the greater part of which is closed under normal circumstances, and open only when the tissue requires a large increase in the supply of blood. In muscle the capacity of this irrigation system may be increased 750 times during activity. Such changes (if of widespread occurrence) will affect arterial pressure by their influence on the total capacity of the vascular system as well as by alterations in the peripheral resistance.

Capillary Fragility. The walls of the capillaries are lined continuously with endothelial cells, which may not normally allow the proteins of the plasma to pass through

---

* LANDIS. *Heart*, 1930, **15**, 209.

When subjected to a sufficient pressure, however, the walls give way, and then blood escapes into the surrounding tissues, causing a local discoloration. When seen in the skin such patches appear as red spots called *petechiæ*. In certain conditions (*v.* Ch. XLII) such as scurvy, these petechial hæmorrhages occur very much more easily than in normal states. The tendency of the skin capillaries to rupture can be found by means of Göthlin's fragility test. A sphygmomanometer armlet is placed round the upper arm, and raised to, and kept at, 50 mm. Hg for fifteen minutes. The armlet is then removed and the skin of the antecubital fossa examined for petechiæ. Not more than eight should be present in a circle of 6 cm. diameter, when examined in a strong light.*

THE CIRCULATION IN THE VEINS

In the veins there must be a constant decrement of pressure as we pass from the periphery towards the heart. Owing to the fact that no appreciable resistance lies between the veins and the heart, the pressure gradient necessary to maintain a constant flow through these vessels is very small. Thus, in the horizontal posture, the pressure in the femoral veins may be from 5 to 10 mm. Hg, and in the external jugular vein from 0 to — 2 mm. It is convenient for clinical purposes to observe the extent to which the external jugular vein is filled with blood when the patient lies down with the head raised. The vein then acts as a manometer, and is seen to be distended to a level which is normally 0 to 2 cm. below a horizontal line passing through the junction of the sternum with the manubrium (Fig. 438). The pressure in the great veins near the heart is generally below that of the atmosphere, owing to the aspiration of the thorax, and becomes still lower during inspiration. Opening the thorax, therefore, causes a rise of pressure in all the large veins. In the latter, the pressure is lowered by vigorous action of the heart pump and raised when this fails in any way. In the peripheral veins the pressure is more dependent on the flow through the corresponding arteries. If an artery of a limb be ligatured, the pressure in the small veins of the limb sinks until it is reduced to the pressure in the nearest large trunk in which a flow of blood continues.

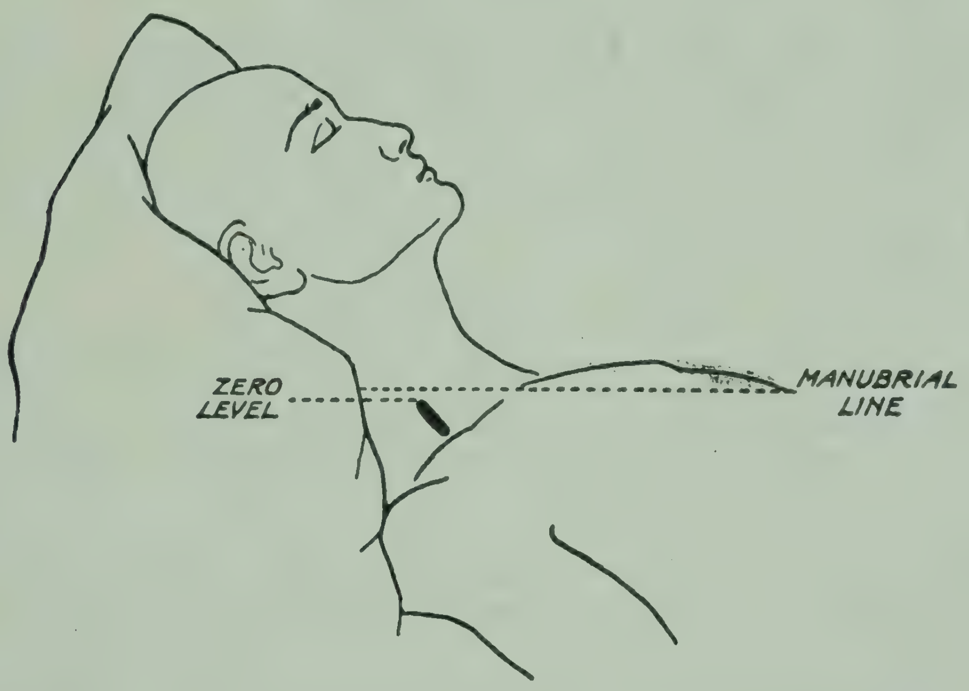


FIG. 438. Level of filling of external jugular vein. (LEWIS.)

If a vein is ligatured, the blood usually finds its way by anastomotic channels into another vein. But if the occluded vein is large, one result will be a rise of capillary and venous pressures distal to the occlusion, and this will continue until adequate anastomoses have been opened up. When all the veins leading from a part, such as a limb, are compressed, the capillary and venous pressures behind the occlusion will rise until they either overcome the resistance, or finally equal the arterial pressure. Thus, if a sphygmomanometer armlet be placed on the upper arm, and inflated to 50 mm. Hg, the venous pressure in the arm veins will rise until it exceeds 50 mm. Hg, so that all the veins are engorged, and if the pressure be raised to 70 mm. the venous pressure will rise accordingly, and so on.

---

* BELL, LAZARUS and MUNRO. *Lancet*, 1940, **239**, i, 154.

The Venous Pulse. Each cardiac cycle causes variations in the pressure in the great veins next the heart in two ways :

(1) By the transmission backwards along the veins of the alterations in the intra-auricular pressure.

(2) By the diminution in the volume of the heart in consequence of the expulsion of its blood along the arteries with each heart beat.

On this account the jugular veins show with each heart beat pulsations which are complex in character and resemble those occurring in the auricle (*vide* p. 561). They can be recorded by the use of the polygraph (Fig. 432, p. 632), by applying an open capsule over the lower part of the external jugular vein and leading to the recording tambour. Jugular tracings (Fig. 439 and 367, p. 564) give information as to the time at which auricular systole occurs, as compared with ventricular systole, which can be deduced from the arterial tracing. It will be seen that each heart beat gives rise to three positive waves of pressure within the veins. These three undulations are analogous to those given in Fig. 366, p. 561, as occurring in the auricular tracing. We regard *a* as due to auricular contraction, *c* as the elevation due to closure of the auriculo-ventricular valves, *v* as the elevation due to the accumulation of blood in the auricles during the ventricular systole ; the crest of *v* marks the end of ventricular systole. The curve *c* is probably due in part also to an impact transmitted from the underlying carotid artery. These venous pulsations are much more marked in cases of heart disease, where there is overfilling of the venous system, often combined with incompetence of the tricuspid valve.

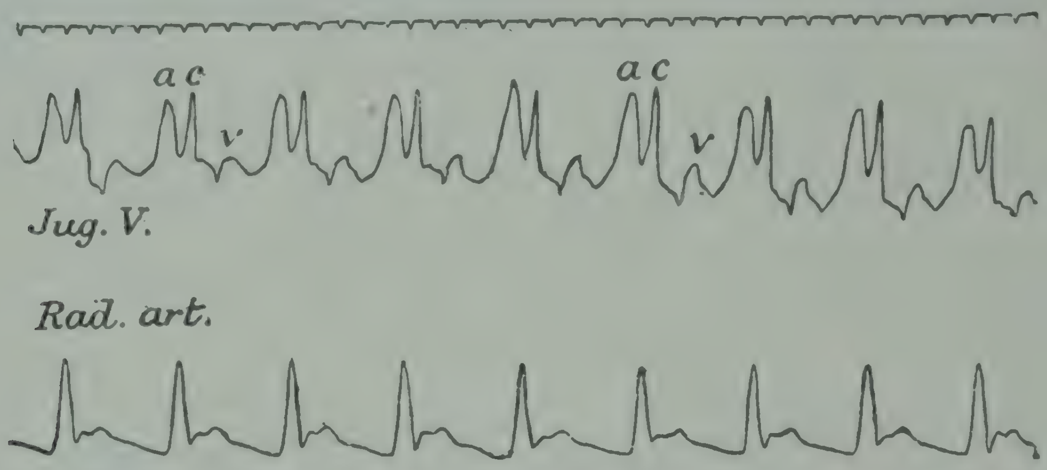


FIG. 439. Venous Pulse Tracing from Jugular Vein compared with the Arterial Pulse Tracing from the Radial Artery.

In cases of heart block (p. 554) the auricle and ventricle contract independently (complete heart block) or at a ratio of two or more auricular contractions to one ventricular. These altered conditions are shown on the jugular and arterial pulse tracings in which jugular tracings show the auricular contractions and the arterial tracings the ventricular contractions (Fig. 358, p. 554).

Valves. Besides the favourable influences exercised on the circulation through the veins by the aspiration of the thorax, a considerable part is played in the venous circulation by the emptying of the veins by outside pressure upon their walls. The adjuvant effect of muscular movement on the circulation through the veins is rendered possible by the existence in these vessels of valves, which are semi-lunar folds of the intima projecting into their lumen, and so arranged that they allow the passage of blood only towards the heart. Two such folds are, as a rule, situated opposite to each other. Every movement of a limb, active or passive, causes an external pressure on the veins and empties them towards the heart. Thus in walking, each time the thigh is moved backwards the femoral vein becomes empty and collapses, and fills again as soon as the leg is brought forward to its former position. When muscular movements become general, as in walking or running, the active compression of the veins thus brought about plays an important part in hurrying the blood into the right heart, so that the output of this organ is increased.

Effect of Gravity. Since the blood in the vessels is subject to the influence of gravity, we might expect to find that, in the erect posture, the pressure in the veins of the foot was equal to the pressure in the veins at the level of the heart *plus* the pressure of the column of blood between these veins and the heart, *i.e.* about a metre of blood, but it is nearly always less than this, and if the leg is exercised, may be little higher than that in the veins of the hand. When a man assumes the upright position, the arteries of the leg and foot contract until, under the combined influence of the heart's contraction and gravity, the blood supply to the capillaries is just adequate. The return of the blood from the dependent parts cannot be ascribed entirely to the negative intrathoracic pressure, but is largely due to the contractions of the muscles of the limb, which press on all the deep and superficial veins, and in virtue of the valves force the blood contained therein past Poupart's ligament into the abdomen. This explains why it is so difficult to stand for any length of time without moving, and emphasises the need of moderate exercise for the maintenance of a normal circulation.

For Reference

EYSTER (1926). Venous Pressure and its Clinical Applications. *Physiol. Rev.*, **6**, 281.
FRANKLIN (1937). A Monograph on Veins. (London : Baillière, Tindall and Cox.)
FRANKLIN (1928). The Physiology and Pharmacology of the Veins. *Physiol. Rev.*, **8**, 346.
HARRISON (1938). Arterial and Venous Pressure Factors in Circulatory Failure. *Physiol. Rev.*, **18**, 86.

CHAPTER XXXIV

LOCAL REGULATION OF BLOOD FLOW

THE NERVOUS CONTROL OF THE BLOOD VESSELS

EVERY organ of the body requires an increased blood supply during activity, which means that blood must be diverted from the inactive to the active tissues. This is chiefly effected through the central nervous system, though local and chemical mechanisms also play some part in the process.

The Vascular Nerves. Our knowledge of the nervous control of the blood vessels dates from the discovery, by Claude Bernard,* that if, in the rabbit, the cervical sympathetic on one side be divided, the vessels in the corresponding ear dilate. Vessels previously invisible come into prominence, and on account of the greater flow of blood, the ear on the side of the section becomes warmer than the normal ear. If the head end of the divided sympathetic nerve be stimulated, all the vessels of the ear contract, and the ear becomes colder than the other. The fact that the dilatation of the vessels produced by section lasts for months shows that the ear vessels are normally under the influence of "tonic," *i.e. continual*, constrictor impulses proceeding to them along the cervical sympathetic.

The paralysis of the ear vessels, though lessening the local resistance to the flow of blood, affects too small a vascular area to have any marked influence on the general arterial pressure. But if the spinal cord be divided at the first dorsal nerve, a wide area is deprived of vaso-constrictor control and the arterial pressure sinks considerably. The heart after the section beats more rapidly than before, so that the fall of pressure must be ascribed to a lowering of the resistance to the flow of blood.

The source of these impulses may be made out by studying the effect of sections through different levels of the nervous system. Division of the cord at, or below, the second lumbar nerve causes no effect on the blood pressure. On making a section above the first lumbar nerve, the effect produced increases progressively until the first dorsal roots are reached, where it is maximal; stimulation of the lower end of the cut dorsal cord causes widespread vascular constriction and a large rise of blood pressure.

We must therefore conclude that all the vessels in the body are kept in a state of tonic contraction by impulses arising from a higher centre, and then passing out of the cord by the dorsal and upper lumbar nerves.† This conclusion is confirmed by the fact that, whereas stimulation of the ventral roots of the cervical, lower lumbar and sacral nerves has no influence on the blood pressure, a rise of arterial pressure can be obtained by stimulating any of the ventral roots from the first or second dorsal to the second or third lumbar. The same effect is produced by stimulation of the white rami communicantes from these roots to the sympathetic system, and by excitation of the sympathetic chain itself, or of the splanchnic nerve.

Since the plain muscle of the blood vessels exhibits an automatic tonus, complete nervous control of these tubes can be secured only by the provision of two sets of nerves: one set—constrictor—which will increase the state of constriction; another set—dilator—which will diminish the tone of the arteriole muscle, and cause vascular dilatation. Our knowledge of the existence of this second class of nerve fibres to the vessels we owe also to Claude Bernard,* who observed that stimulation of the chorda tympani nerve increased the blood flow through the sub-maxillary glands.

---

* CL. BERNARD. *C. R. Soc. Biol.*, 1852, **4**, 168.
† GASKELL. *J. Physiol.*, 1885, **7**, 1.

Criteria of Vasomotion. In order to investigate the state of tone of the small blood vessels it is necessary to have means for investigating the condition of the blood flow through different parts of the body. It is obvious that if the arterioles A in the organ B dilate (Fig. 440), the first effect is a diminution of the resistance to the flow of blood into the capillaries beyond. If the arterial pressure in the trunk C remains constant, a local diminution of resistance in A will at once determine an increased flow of blood through the arterioles. If the organ is distensible, the increased pressure in the arterioles and capillaries will cause a swelling of the whole organ. But the same effect on intracapillary pressure and therefore on the volume of the part, may also be caused by obstruction to the flow of blood from the veins V. It is only when, the arterial pressure in C remaining constant, dilatation of an organ is *accompanied by an increased blood flow through it* that it may be taken as an expression of vasodilatation in the arteries with which it is supplied.

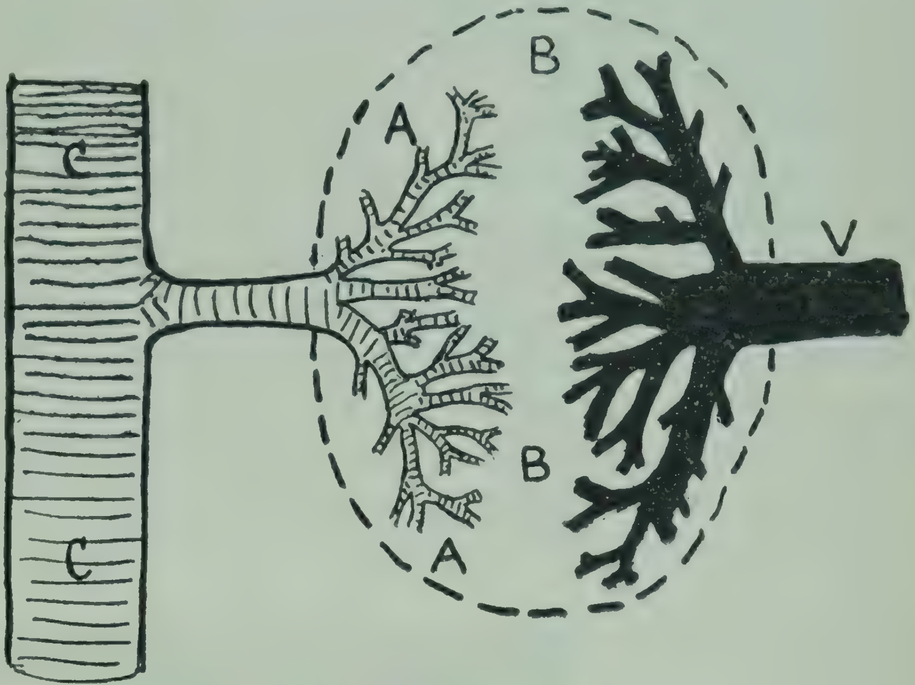


FIG. 440.

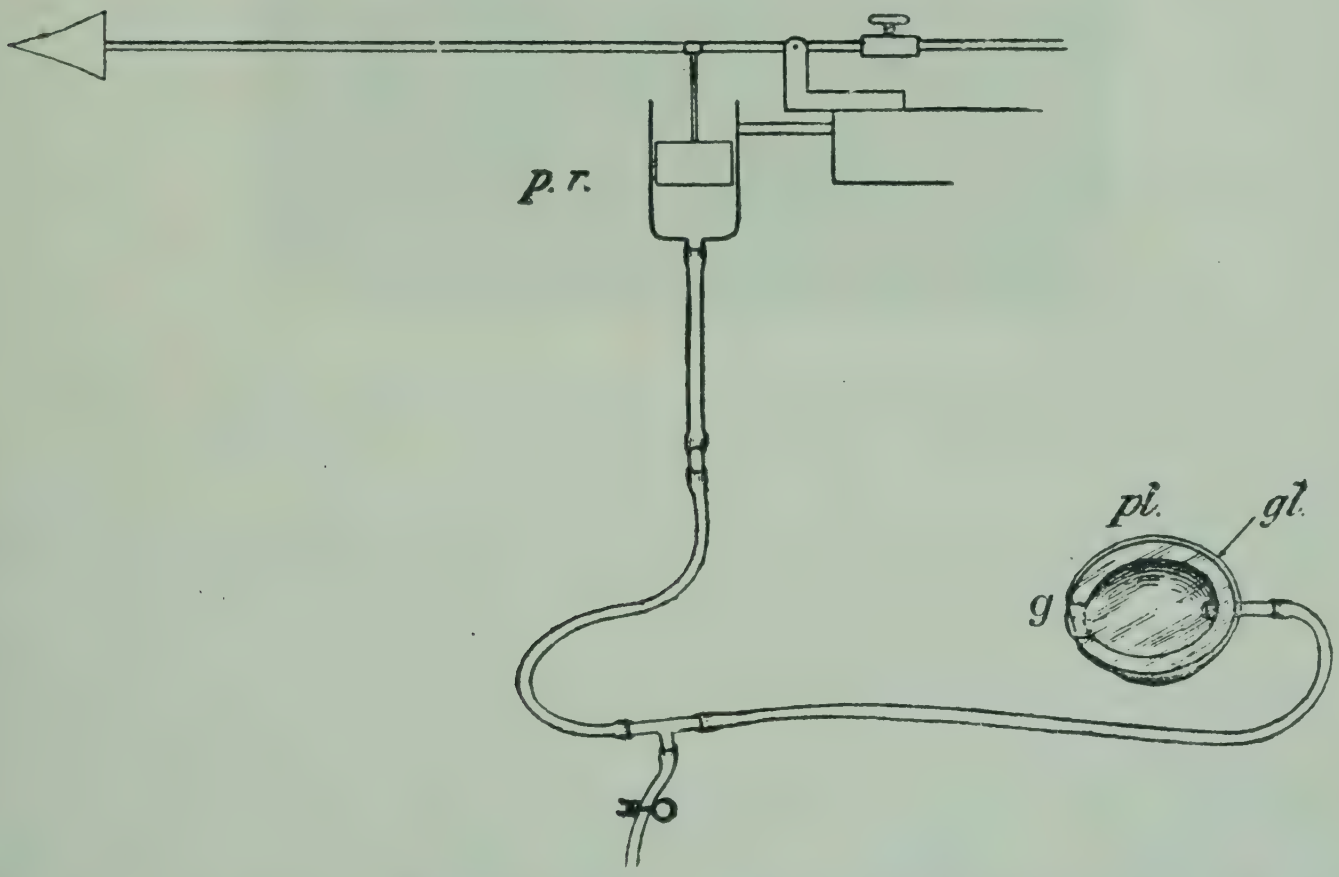


FIG. 441. Diagram of Plethysmograph. *pl.*, the plethysmograph, covered by a glass plate, *gl*; the vessels pass in through the groove *g*; *p.r.*, the piston recorder. (From SHARPEY SCHAFER'S "Experimental Physiology," Longmans, Green.)

We can therefore use the following criteria for the occurrence of a vasodilatation (and their converses for vasoconstriction) in the arterial supply to any part or organ :

---

* BERNARD. *C.R.*, 1858, **46**, 159.

(1) If the surface of the part is translucent, the increased filling of the capillaries will cause redness.

(2) The increased size of the vessels will cause an *increase in the volume* of the organ concerned.

(3) An increased velocity of blood flow will usually raise its temperature.

(4) An increased velocity in the arteries going to the part may be interpreted as due to vasodilatation.

(5) The increased flow may also be detected by measuring the total volume of blood which flows from the vein in a given time.

Of these methods the two most used are those based on determination both of the volume of the part, and of either the arterial inflow to, or of the venous outflow from, the part. A fallacy may arise, however, unless means be taken to ensure that the general arterial pressure remains constant during the experiment. A rise of arterial pressure will cause a *passive* expansion of the vessels and of the part supplied, and also increased velocity of blood flow through the part. In all cases, therefore, it is necessary to record the arterial pressure. We may take as an instance an experiment on the blood supply to the kidney.

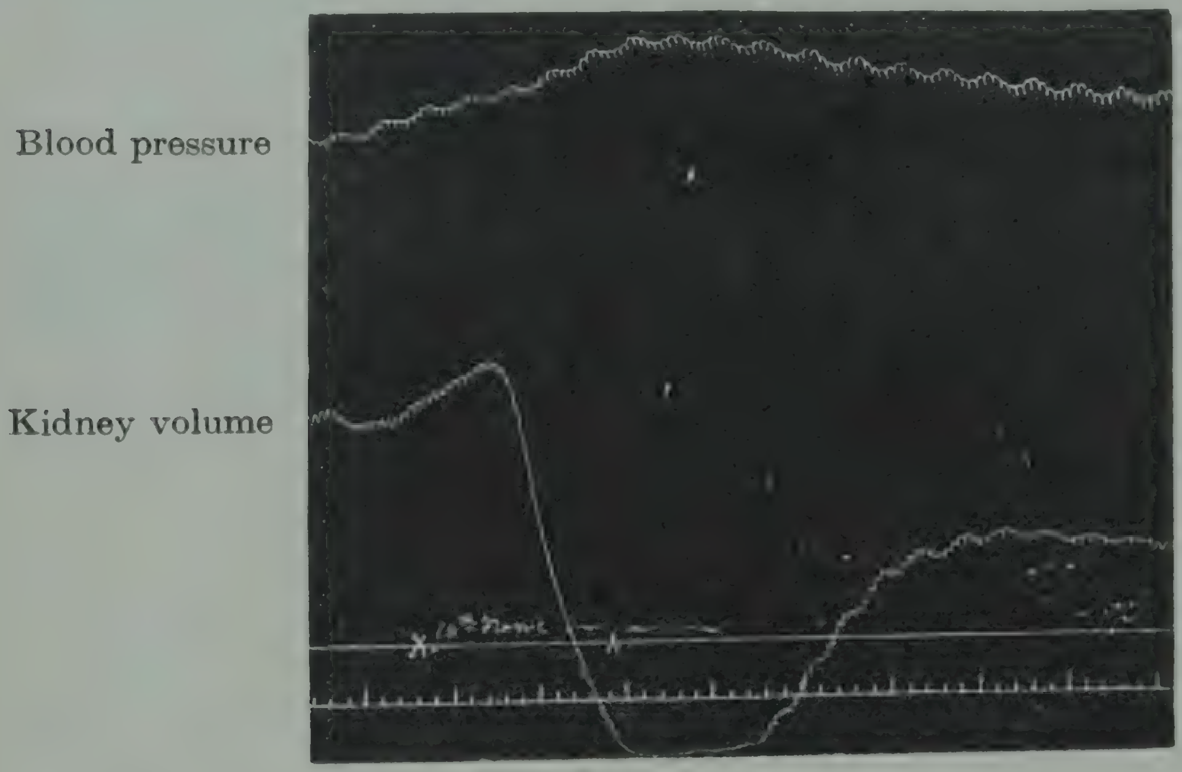


Fig. 442. Simultaneous Tracings of Carotid Blood Pressure and Volume of Kidney. Between × and × the peripheral end of the divided tenth dorsal nerve was stimulated. Time-marking = seconds. (Bradford.)

For this purpose we may determine the change in the volume of the organ by the use of a *plethysmograph*,* (Fig. 441); this can be made of Stent's wax previously moulded to the size of the organ whose volume is the object of investigation. In one side of the box a depression is left, sufficient to accommodate easily the vessels, nerves (and ureter) going to the kidney. The plethysmograph is covered with a glass lid which is made air-tight by means of vaseline, the space between the lid and the vessels being also packed with cotton-wool and vaseline. A tube from the plethysmograph leads to a piston recorder. Every variation in the volume of the organ causes a movement of air into or out of the plethysmograph and thus gives rise to a corresponding movement of the recording lever.

The kidney being placed in some such apparatus, and the arterial pressure being recorded, as in Fig. 442, the peripheral end of the anterior root of the tenth dorsal nerve was stimulated ; this was followed by a rise of blood pressure together with a diminution of the kidney volume. The increased blood pressure would by itself tend to force more blood into the kidney. Since the kidney volume diminished there must have been active contraction of the arterioles of the kidney. This contraction of the vessels must have taken some part at any rate in the rise in general blood pressure. If the plethysmograph had been used alone, it would have been impossible to determine whether the shrinkage of the kidney might not be due to a lowering of arterial pressure. On the other hand, without the volume record it would only have been possible to infer that there was increased peripheral resistance somewhere or other in the body.

---

* Schäfer and Moore, *J. Physiol.*, 1896, **20**, 1.

Instead of taking the volume of the kidney, it would now be more usual to measure the blood flow through its vessels; this would have been found to be diminished by the stimulation. The method is of great value in dealing with the circulation through small organs, such as the submaxillary glands. In such a case it is usual to prevent the clotting of the blood by the preliminary injection of heparin, and then, after placing a cannula in the efferent vein of the organ, to allow blood from the cannula to flow into an automatic measuring apparatus, *e.g.* that of Gaddum.*

For recording the arterial inflow the thermostromuhr is now often employed (p. 624).

Course of the Vasoconstrictor Fibres. The vasoconstrictor fibres belong to the sympathetic system.

The position of the synapses on the path of the fibres is shown by the use of nicotine.† After injection of nicotine, stimulation of the pre-ganglionic fibres is without effect, although the normal results follow stimulation of the post-ganglionic non-medullated fibres.

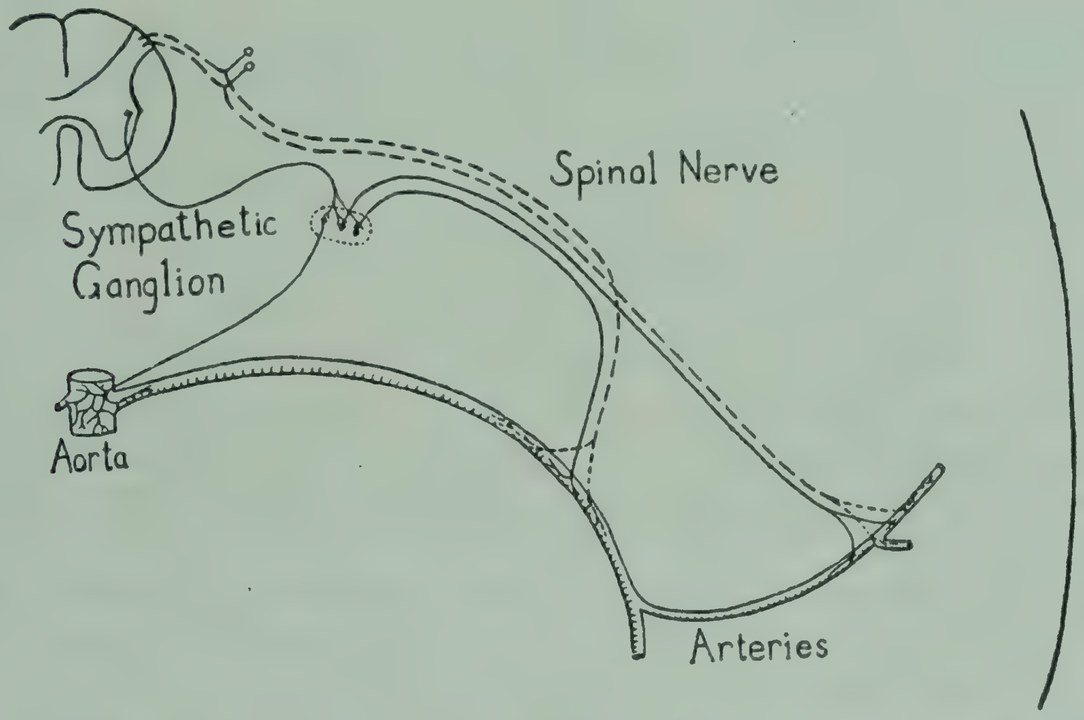


FIG. 443. Diagram showing the distribution of nerves to the arterial vessels. Sympathetic fibres continuous lines. Spinal afferent fibres dotted lines. (RANSON. *Ann. Int. Med.*, 1933, **6**, 1013.)

By this method it has been determined that all the vasoconstrictor nerves of the body leave the spinal cord by the ventral roots of the spinal nerves from the first thoracic to the third or fourth lumbar inclusive. From the roots they pass by the white rami communicantes to the ganglia of the sympathetic chain. Here they take different courses according to their destination.

The vasoconstrictor nerves to the periphery of the body (Fig. 443), including the skin and muscles, pass by the grey rami communicantes to the mixed spinal nerves and are distributed with them. The distribution is strictly unilateral,‡ stopping abruptly at the mid-line, in spite of the fact that the blood vessels themselves anastomose freely across the mid-line. It would seem therefore that, so far as the skin supply is concerned, the peri-arterial plexus described by Woollard § is not concerned with the transmission to a distance of vasoconstrictor impulses.

---

* GADDUM. *J. Physiol.*, 1929, **67**, 1.P.
† LANGLEY. *J. Physiol.*, 1890, **11**, 123.
‡ GILDING. *J. Physiol.*, 1932, **74**, 34.
§ WOOLLARD. *Heart*, 1926, **13**, 319.

The fibres to the head and neck leave by the first four thoracic nerves, and pass to the superior cervical ganglion. Here they end, and the impulses are carried by a fresh relay as non-medullated fibres on the walls of the carotid artery and its branches.

The constrictors to the fore limb in the dog leave the cord by the fourth to the tenth thoracic nerves. The fibres run to the stellate ganglion, where they all end in *synapses*. The impulses are carried on by the grey rami to the nerves which make up the brachial plexus, and run down in the branches of this to the vessels of the fore limb.

The constrictor impulses to the hind limb in the dog arise from the nerve roots between the eleventh thoracic and third lumbar roots, end in connection with cells in the sixth and seventh lumbar and first and second sacral ganglia of the sympathetic chain, whence the impulses are carried by grey rami to the nerves making up the sacral plexus.

The chief vasoconstrictor supply to the abdominal viscera is the *splanchnic nerve*. This receives fibres of the white rami from the lower seven thoracic and upper two or three lumbar roots, the latter fibres often taking a separate course as the lesser splanchnics. The fibres pass through the sympathetic chain of the thorax without interruption, and for the most part have their cell stations in the collateral ganglia, especially the cœliac ganglion, whence a thick meshwork of non-medullated fibres is distributed along all the vessels of the abdominal viscera. The area of the vessels innervated by this nerve is so large that section of it on both sides causes a definite fall in the arterial pressure.

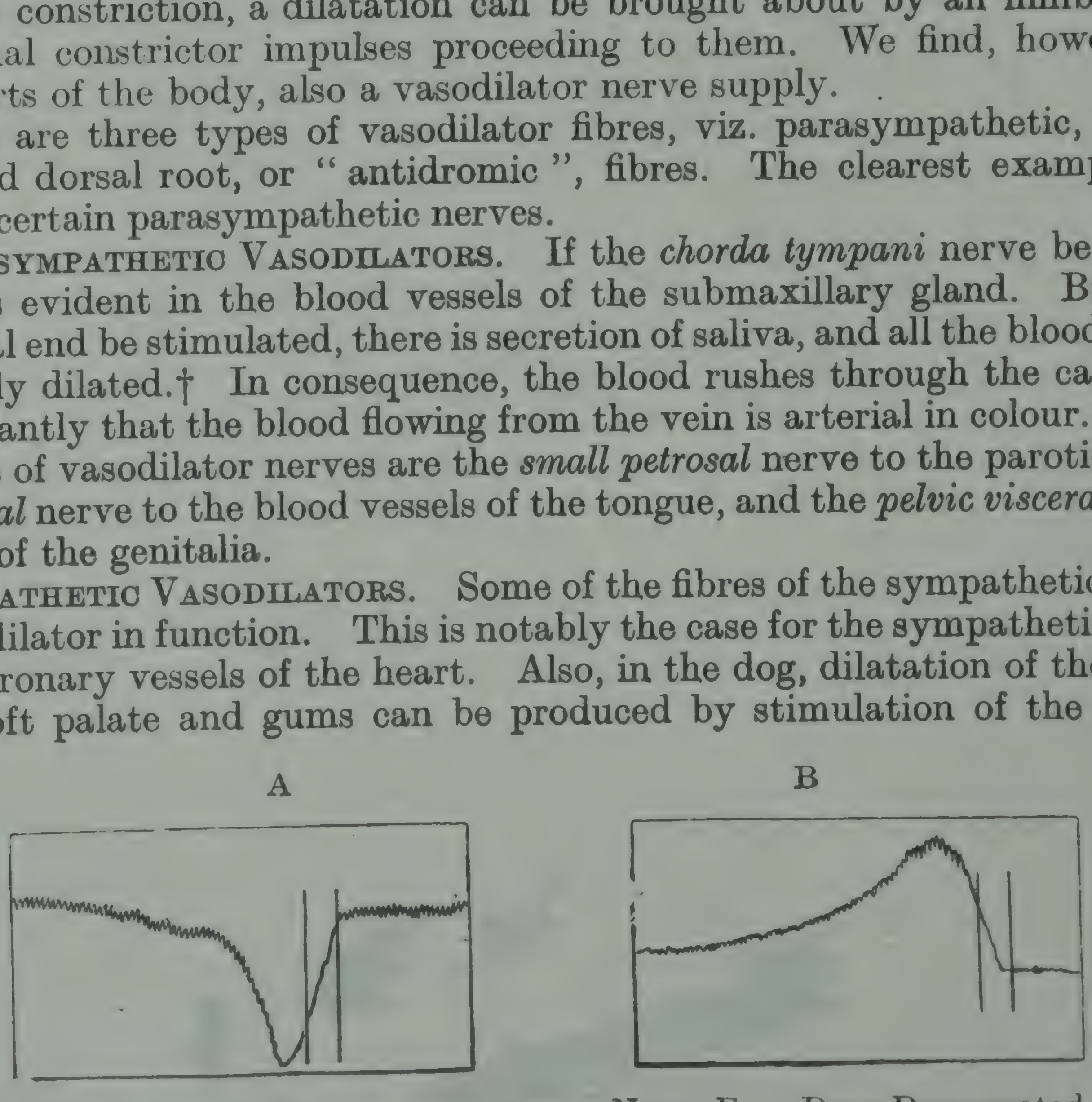
CONSTRICTORS TO VEINS. Often the veins are also supplied with sympathetic vasoconstrictor fibres *; their functions are less conspicuous than those of the arteries, but must be borne in mind, particularly in the interpretation of plethysmographic tracings.

Vasodilator Nerves. Since the arteries are usually in a condition of moderate constriction, a dilatation can be brought about by an inhibition of the normal constrictor impulses proceeding to them. We find, however, in many parts of the body, also a vasodilator nerve supply.

There are three types of vasodilator fibres, viz. parasympathetic, sympathetic and dorsal root, or "antidromic", fibres. The clearest examples are found in certain parasympathetic nerves.

PARASYMPATHETIC VASODILATORS. If the *chorda tympani* nerve be cut, no change is evident in the blood vessels of the submaxillary gland. But if its peripheral end be stimulated, there is secretion of saliva, and all the blood vessels are largely dilated.† In consequence, the blood rushes through the capillaries so abundantly that the blood flowing from the vein is arterial in colour. Other examples of vasodilator nerves are the *small petrosal* nerve to the parotid gland, the *lingual* nerve to the blood vessels of the tongue, and the *pelvic visceral* nerves to those of the genitalia.

SYMPATHETIC VASODILATORS. Some of the fibres of the sympathetic system are vasodilator in function. This is notably the case for the sympathetic supply to the coronary vessels of the heart. Also, in the dog, dilatation of the vessels of the soft palate and gums can be produced by stimulation of the cervical



Nerve Freshly Divided
Constriction.

Nerve Four Days Degenerated.
Dilatation.

FIG. 444. Plethysmographic Tracing of Hind Limbs, showing effect of Stimulating the Sciatic Nerve on the Volume of the Limb, A, immediately after section of the nerve; B, four days after section. The nerve was stimulated between the two vertical lines. Curves to be read *from right to left*. (BOWDITCH and WARREN.)

---

* HOOKER. *Amer. J. Physiol.*, 1918, **46**, 591; DONEGAN. *J. Physiol.*, 1921, **55**, 22

† CL. BERNARD. *C. R.*, 1858, **46**, 159.

sympathetic of the same side, or of the stellate ganglion, or its *rami communicantes*. Again, in the muscles of the dog, sympathetic dilators preponderate. In other species, the sympathetic supply to muscles is mainly or wholly vasoconstrictor, as in rabbit and monkey. Usually, special conditions must be present in order to render possible the demonstration of sympathetic vasodilators. Thus, stimulation of the splanchnic nerves may cause a fall of blood pressure, provided that the constrictor fibres, which predominate, have been paralysed by the previous administration of large doses of ergotamine,* or of dibenamine.† Again, the presence of vasodilator fibres in the lumbar sympathetic chain supplying the hind limb in the dog can be demonstrated when this is stimulated, provided that some adrenaline is present in the circulating blood.‡ This effect is intensified by eserine and annulled by atropine. Hence it may be concluded that these vasodilators, though sympathetic, are cholinergic: they chiefly supply the muscles. Similar cholinergic vasodilators to muscle have been demonstrated in the cat after paralysis of constrictors by dibenamine.§

VASODILATORS IN SPINAL NERVES. In addition, there are in the mixed spinal nerves, fibres which are vasodilator. Since, however, the mixed nerves to the limbs also contain constrictor fibres, to muscles, but chiefly to the skin, the effect of the constriction overpowers any effects due to dilator fibres. Moreover, since the dilators apparently do not conduct any tonic influences, the only effect of section of a mixed nerve is that due to the removal of the tonic constrictor influences, and the vessels in the area of distribution therefore dilate.

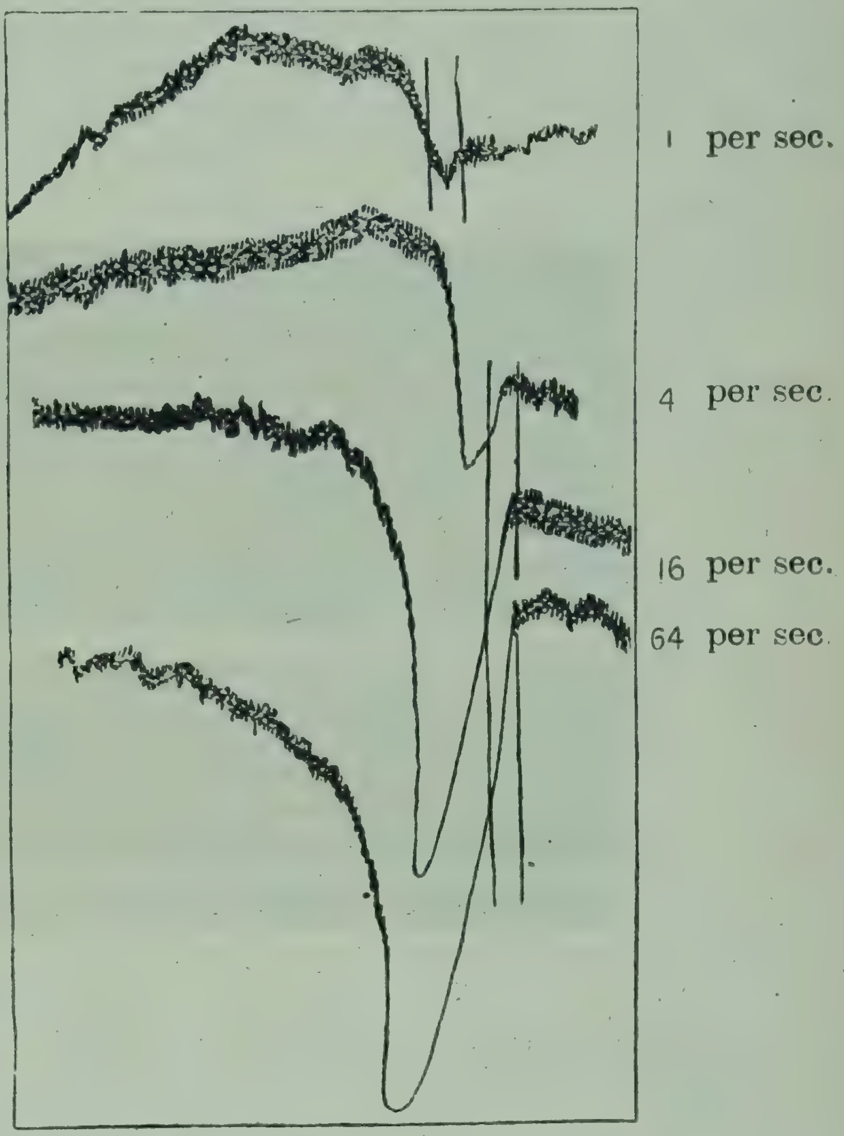


FIG. 445. Effect on the Volume of the Hind Limbs of the Cat of Stimulating the Sciatic Nerve with Induction Shocks at Different Rates. With one shock per second there is dilatation, whereas with 64 shocks per second the effect produced is vasoconstriction. Read *from right to left.* (BOWDITCH and WARREN.)

Various methods have been employed to show the presence of dilator fibres in such a mixed nerve trunk. One method is to remove the skin, which is supplied almost entirely by constrictor fibres, whereby vasodilator responses are more easily revealed.

Sympathetic vasodilators to the hind limb muscles have been shown (cat)|| by slow stimuli (15/sec.) to the sympathetic chain, or by stimulation after paralysis of vasoconstrictors by dibenamine; degeneration of afferent fibres in the chain did not eliminate these affects but degeneration of efferent fibres did; these vasodilators are cholinergic, though sympathetic.

ANTIDROMIC VASODILATORS. Vasodilators to the skin appear to fall into an entirely separate category; some of the muscle dilators may also be

---

* DALE. *J. Physiol.*, 1913, **46**, 294.

† NICKERSON and GOODMAN. *J. Pharmacol.*, 1947, **89**, 167.

‡ BURN. *J. Physiol.*, 1932, **75**, 144.

§ FOLKOW *et al. Acta Physiol. Scand.*, 1948, **15**, 401.

|| FOLKOW and UVNAS. *Acta Physiol. Scand.*, 1948, **15**, 389.

included in this. The presence of vasodilators to the limbs in mixed nerves can be shown in various ways. Two of these depend on the unequal time taken for the constrictor and dilator fibres to degenerate, and on the varying excitability of the two sets of fibres to different kinds of stimulation.* Thus, if the sciatic nerve be cut, and after two or three days the peripheral end of the divided nerve be stimulated, dilatation of the vessels is brought about (Fig. 444). The constrictor fibres degenerate before the dilator fibres, so that only the latter respond to stimulation. On the other hand, it is often possible, even in the freshly cut nerve, to obtain dilatation by stimulating its peripheral end with induction shocks repeated at slow intervals—one to four per second, as shown in Fig. 445.

In the light of the evidence presented above, it would seem probable that some, at all events, of the vasodilators concerned are sympathetic vasodilators to muscle, and therefore emerge from the cord by ventral roots. But some limb dilator fibres are not connected with the sympathetic system, but with the *dorsal* nerve roots. It was shown by Stricker in 1876 that dilatation of the vessels of the hind limb could be produced by stimulating the dorsal roots of the nerves going to the limb, *i.e.* far below the point of origin from the cord of the constrictor fibres to the same part of the body. This has often been confirmed.† Stimulation of the dorsal roots, either before or after they have passed through their ganglia, causes dilatation of the vessels in the area of skin and muscle ‡ supplied by the roots (Fig. 446). This effect is not destroyed by previous section of the dorsal roots on the proximal side of the ganglia, showing that the fibres by means of which the dilatation is produced have the same course as the ordinary sensory nerves to the limbs. Since the impulses on artificial stimulation pass in a direction opposite to that taken by the normal impulses, Bayliss designated them *antidromic* impulses. The phenomenon of a nerve fibre functioning in both directions is without analogy in our knowledge of the other nerve functions of the body, and there is doubt as to the relation of the phenomena obtained by stimulation of the dorsal roots to those of vasodilatation in the normal state, as, *e.g.* to local stimulation by mechanical means or by irritants, such as mustard, applied to the skin.

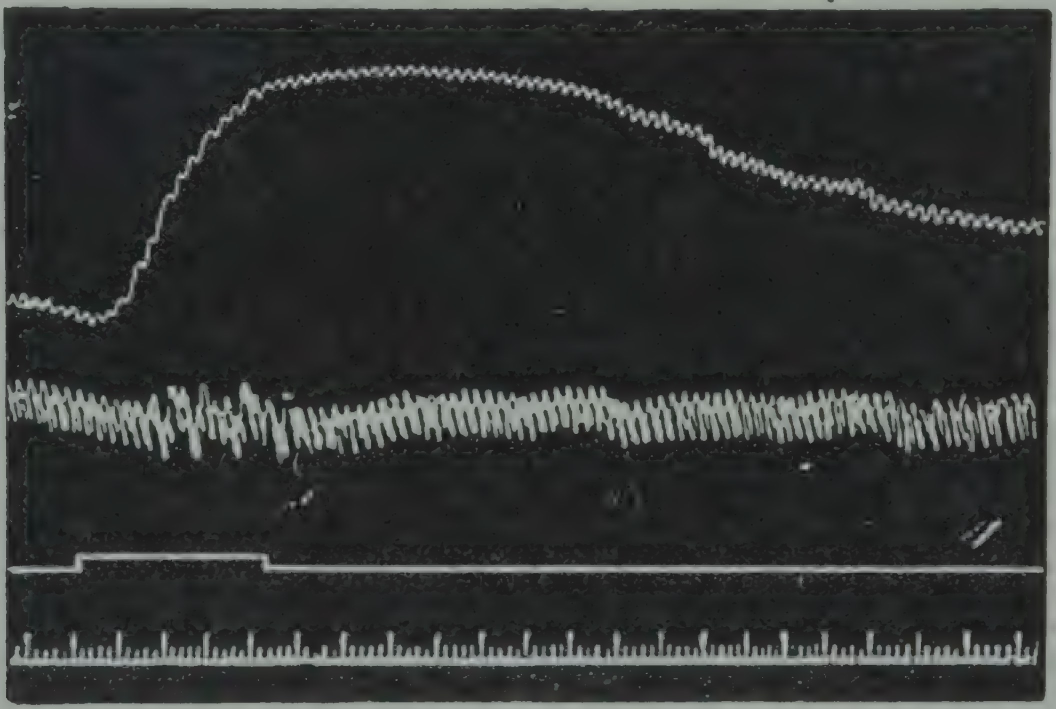


FIG. 446. **Effect of Excitation of Peripheral end of the Seventh Lumbar Dorsal Root in the Dog.** (BAYLISS.)

Uppermost curve, volume of left hind limb; next below arterial blood pressure; the third line marks the period of stimulation; bottom line, time-marking in seconds.

---

* BOWDITCH and WARREN. *J. Physiol.*, 1886, **7**, 416.

† BAYLISS, W. M. *J. Physiol.*, 1901, **26**, 173; 1902, **28**, 276; BACH. *Amer. J. Physiol.*, 1946, **145**, 474.

‡ WYBAUW. *Arch. int. Physiol.*, 1938, **46**, 293.

The commonly accepted view of the normal vasodilator response which is effected through the dorsal root fibres is that it is an axon reflex, stimulation of the skin causing dilatation of its arterioles without the afferent impulse reaching the spinal cord. If the dorsal nerve roots concerned are divided, the vasodilatation resulting from local irritation of the skin, *e.g.* by mustard, still persists, but it disappears if the section is made distal to the dorsal root ganglion and time allowed for the fibres to degenerate. It also disappears if the skin is locally anæsthetised by injection of novocain. Hence it is assumed that the axons of some, presumably afferent, nerves branch, as shown in Fig. 447, one branch going to afferent receptors and another to the muscle cells of cutaneous arterioles. Afferent fibres of the trigeminal are believed to behave similarly.

However they reach the minute vessels, there is some evidence that the results there produced by the nerve impulses may be due to the liberation of potent chemical substances in the tissues of the locality. These may be histamine * or acetyl choline,† the former dilating capillaries, the latter the arterioles.

It may be noted that in certain inflammatory conditions in the dorsal root ganglia, there is reddening of the skin in the areas supplied by the respective roots, often followed by a vesicular eruption (Herpes Zoster) also confined to those areas.

The Vasoconstrictor Centre. The portion of the medulla concerned with the sending out of the tonic vasoconstrictor impulses is spoken of as the vasoconstrictor, or vasomotor, centre. Section of the crura cerebri, or of the brain stem at the upper border of the fourth ventricle, leaves the blood pressure unaffected. Destruction of a small region of the medulla situated on each side of the middle line in the neighbourhood of the facial nucleus, *i.e.* in the forward prolongation of the lateral columns, after they have given off their fibres to the decussating pyramids, causes an immediate and maximal lowering of the blood pressure. Stimulation of a point, the fovea inferior on the floor of the fourth ventricle, and also often of the facial colliculus, causes a rise of blood pressure, while excitation of a point, the obex, near the calamus scriptorius, causes a fall. The former effect is probably due to excitation of the vasoconstrictor centre,

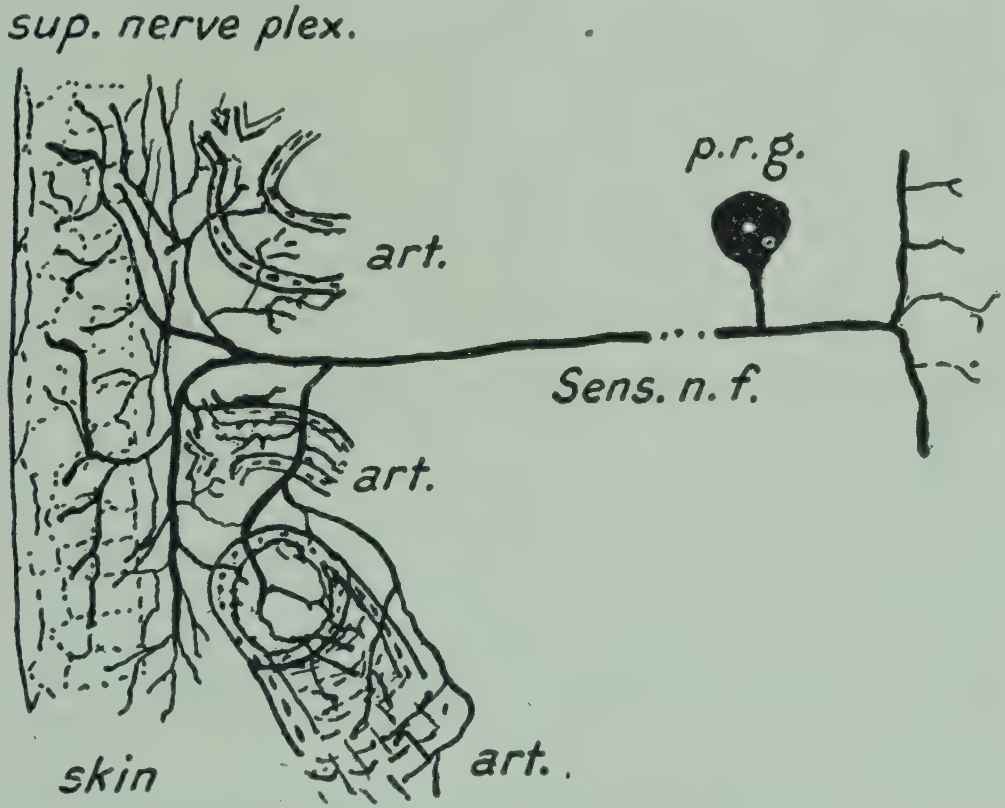


FIG. 447. Diagram to illustrate the Production of Vasodilatation in the Area of Distribution of a Sensory Nerve.

p.r.g. Dorsal root ganglion. *Sens.n.f.* Sensory nerve fibre, branching to supply dilator fibres to the skin arteries, and sensory fibres to the skin.

---

* LEWIS and MARVIN. *Heart,* 1927, **14,** 27.

† DALE and GADDUM. *J. Physiol.,* 1930, **70,** 109.

and the latter has similarly been supposed to be the result of stimulation of a vasodilator centre.* The existence of a true vasodilator centre is doubtful, however, since there is evidence that stimuli applied to the depressor point act on some afferent tract which mediates reflex vasodilatation.†

The vasomotor centre is played upon by impulses from the higher centres of the brain, but especially by afferent impulses travelling from all parts of the body, including those carried by the vagi and aortic nerves and by the carotid sinus nerve. The last two produce an inhibitory effect on the centre, so that denervation of the carotid sinuses and section of the aortic nerves results in a rise of arterial pressure. Whether, in the absence of all afferent stimuli, the centre would be active is doubtful; all we know is that the sum of the stimuli acting upon the centre produces a state of average continued activity, which is responsible for the maintenance of arterial tone.

CHEMICAL EFFECTS ON CENTRE. The most important factor for the maintenance of the tonic activity of the centre is the presence of carbon dioxide in the blood. If this be lowered by forcible ventilation of the lungs, so as to remove much carbon dioxide out of the blood, a profound fall of blood pressure occurs owing to paralysis of the vasomotor centre. On the other hand, anything which interferes with the gaseous exchanges of the centre, whether obstruction to respiration, or a failure of the blood supply, as by ligature of the cerebral arteries, calls forth an increased state of activity of the centre. This can be illustrated by observing the changes in the blood pressure produced in a curarised animal by the cessation of artificial respiration.

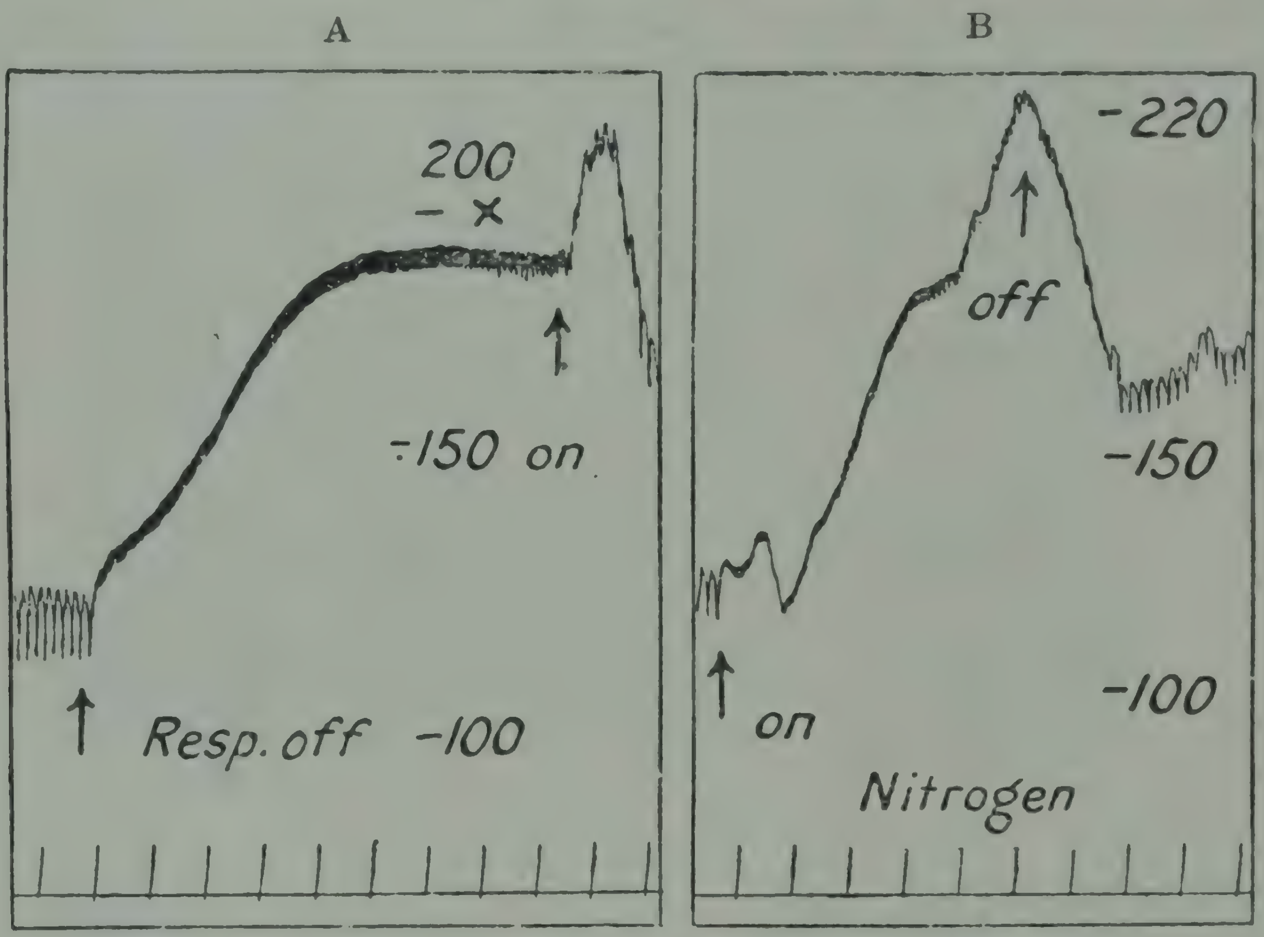


FIG. 448. Blood Pressure Changes in a Cat. (MATHISON.)

A. After cessation of respiratory movements. B. As a result of artificial respiration with nitrogen.

These changes depend partly on the stimulation of the vasomotor and cardio-inhibitory centres by the venous blood, partly on effects upon the carotid region, (*v.* p. 595), and partly on an action upon the heart itself.‡ We will first consider them with both vagi cut. For several seconds after leaving off the artificial respiration the only change noticed is the absence of the respiratory oscillations. Then the blood pressure suddenly rises rapidly (Fig. 448A) and in another ten seconds may reach a height twice as great as it was previously. The heart frequency is almost unaltered. The blood pressure remains at this height for about a minute and then gradually falls, the heart beats becoming smaller and smaller until the pressure has sunk almost to zero. *This fall in pressure is due to the failure of the heart.* The heart gets overfilled, heart block ensues, and it gradually fails. The vessels, however, remain constricted until the death of the animal. This is shown by two facts. If, while the pressure is sinking, artificial respiration be recommenced, the heart, again supplied with oxygen, at once begins to beat more forcibly, and the blood pressure may rise to an even greater height than immediately after the commencement of the asphyxia. Again, if the volume of the kidney be recorded by means of the oncometer, the rise of general blood pressure produced by asphyxia is seen to be accompanied by a rapid shrinking of the kidney, and this shrinking endures until just before the animal dies.

---

* RANSON and BILLINGSLEY. *Amer. J. Physiol.*, 1916, **41**, 85.

† SCOTT and ROBERTS. *J. Physiol.*, 1923, **58**, 168; SCOTT, *J. Physiol.*, 1924, **59**, 443.

‡ MATHISON. *J. Physiol.*, 1910, **41**, 416; 1911, **42**, 283.

Similar results are obtained when the vessels to the brain are ligatured.

We have spoken above of the phenomena of asphyxia as being due to the circulation of venous blood. This, however, comprises two factors, viz. a diminished percentage of oxygen and an increased percentage of carbon dioxide; and the question arises whether one or both of these factors are concerned. It is easy to investigate the action of each separately. A pure oxygen lack may be brought about by allowing an animal to breathe some inert gas, such as nitrogen, or by injecting small amounts of cyanide into the blood;* while the effects of accumulation of carbon dioxide in the blood and tissues may be produced by the administration of mixtures containing excess of oxygen, with varying percentages of carbon dioxide. In the first case, not only will the oxygen supply be reduced, but the tension of the carbon dioxide in the blood will also be kept below normal; in the second case, while the CO_2 tension is raised, the tension of oxygen in the blood will be kept above normal. Mathison found that both conditions may concur in the production of the rise of blood pressure in asphyxia as shown by Figs. 448B and 449. There are certain minor details in these curves which are of interest. When the oxygen of the lungs is rapidly washed out with a neutral gas, the rise comes on about half a minute later than it would with pure asphyxia. In the latter case it seems that the first rise is due to the accumulation of carbon dioxide. The rise under nitrogen, however, when it occurs, is extremely abrupt, and the subsequent fall of blood pressure, *i.e.* the heart failure, is earlier in onset and more rapid than with ordinary asphyxia. When excess of carbon dioxide is administered, *i.e.* 5 to 10 per cent., a large rise of pressure occurs which is almost entirely due to stimulation of the vasomotor centres and resulting constriction of the peripheral arterioles. If a loop of intestine be placed in a plethysmograph, it will be seen that the rise of pressure coincides with a vascular constriction (Fig. 449). The rise of blood

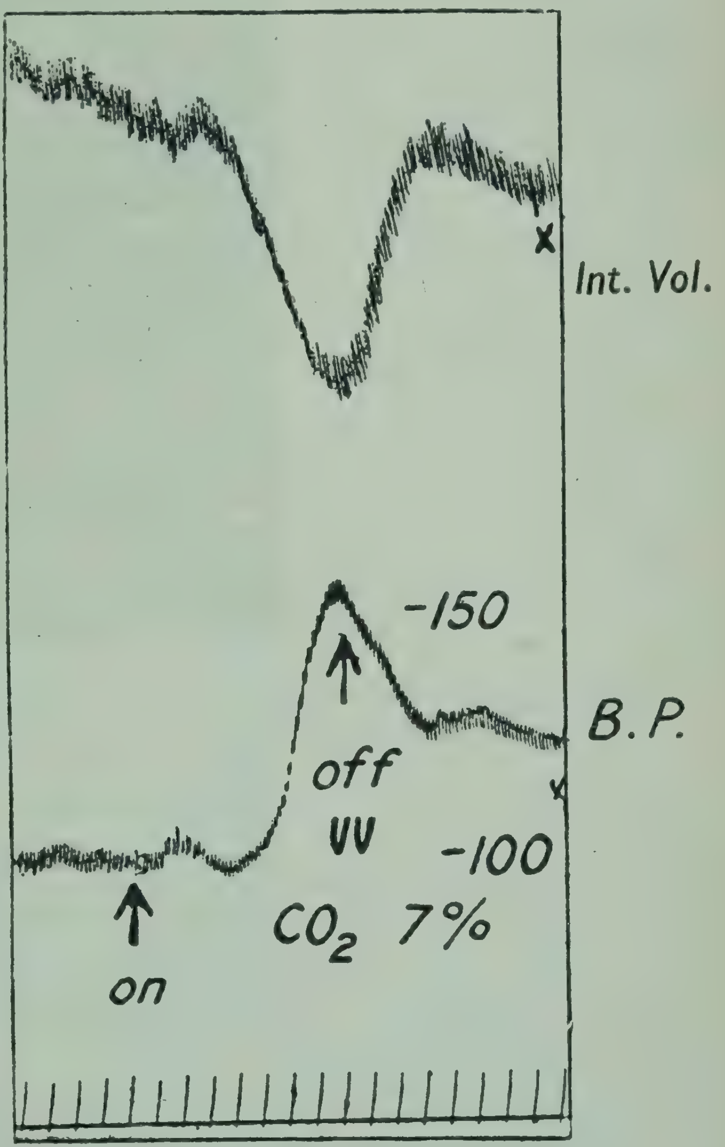


FIG. 449. Tracing of Arterial Blood Pressure and of Intestinal Volume, to show the Influence of a Moderate Increase in the CO_2 tension of the Blood. (MATHISON.)

---

* LOVATT EVANS. *J. Physiol.*, 1919, 53. 17.

pressure due to the vascular constriction may be maintained for a considerable period since we do not get the failure of the heart that is observed in an ordinary asphyxia tracing. If partial oxygen lack, or abnormally increased tension of carbon dioxide, be continued for some time, a state of narcosis or paralysis finally ensues, which affects not only the higher centres but also those of the medulla, so that death may result without convulsions or excessive rise of blood pressure.

The action of cyanide resembles that of nitrogen breathing in its main features, but differs from it in having a more rapid onset, and a slower recovery. This is shown by Fig. 450, which also illustrates another feature of asphyxia. The blood pressure curve shows a second rise ; this is abolished if the suprarenals are removed : it is due to the liberation of adrenaline into the blood stream, which constantly happens in oxygen lack.

Is there any common factor in the two conditions of oxygen lack and carbon dioxide excess, which may account for the similarity in their effects ? It has been shown that, whenever there is a deficiency of oxygen, lactic acid is formed. Lactic acid introduced into the blood stream also causes a rise of blood pressure similar to the asphyxial rise. When the cells of the vasomotor centre are deprived of oxygen, through lack of sufficient oxygenated blood, or are prevented from using it, as by cyanides, it is probable that they themselves produce lactic acid, which, like carbon dioxide, raises the H-ion concentration of their contents and thus excites them.

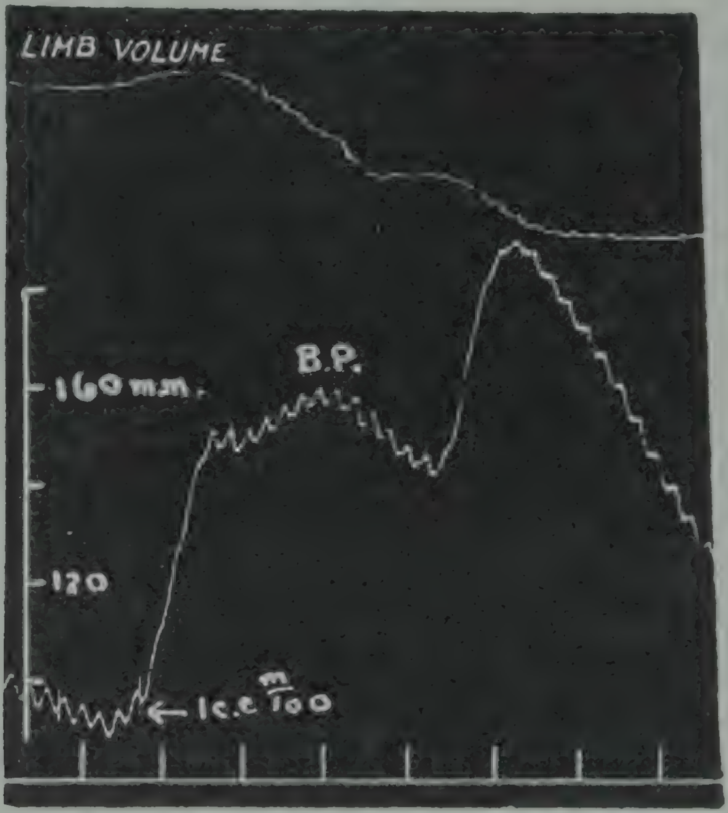


FIG. 450. Decerebrate Cat. Vagi divided. Upper tracing = limb volume. Second tracing = arterial pressure. Lower tracing = time (10 secs.). Injection of 1 c.c. of $\frac{M}{100}$ NaCN into femoral vein. (C.L.E.)

If the vagi be left intact, the blood pressure tracing during asphyxia has quite another appearance. In this case, there is at first only a slight rise of pressure, but the heart begins to beat very slowly. This slow beat is due to excitation of the vagus centre, and is at once abolished by section of the two vagi. The sparing of the heart by means of this vagus action enables it to last longer, so that heart failure comes on rather later than when the vagi are divided. In the increased vagus action, which occurs during asphyxia, three factors are probably involved. The cardio-inhibitory centre in the medulla partakes of the general excitation of the medullary centres, due in the first place to carbon dioxide excess, in the second to oxygen lack. Thirdly, the carotid sinus and aorta are stretched by the rising pressure, and further rise is prevented owing to this : it also causes reflex cardiac inhibition.

RHYTHMIC ACTIVITY. When the arterial pressure is high, especially in asphyxia, waves are often observed on the blood-pressure curve. These are of two kinds. In animals immobilised by curare we may observe oscillations of blood pressure corresponding in frequency with the respiratory rhythm before curarisation. They are due to irradiation of impulses from the respiratory centre to the vasomotor centre. These are the *Traube-Hering waves.** Besides these, others are occasionally seen which arise in a slow rhythmic variation of the constrictor impulses. These waves are known as the *Mayer waves* and are much slower in their rhythm than the respiratory waves. They are observed especially in dogs poisoned with morphia, or after hæmorrhage (Fig. 451). The presence of these waves may be ascribed to a state of rhythmically fluctuating excitability of the vasomotor centre. This excitation may arise in various ways. Often the rhythmicity is related to a rhythmically waxing and waning excitability of the respiratory centre.*

---

* HERING. *Wiener Sitzungsber.*, 1869, **60**, 829 ; TRAUBE. *Z.f.d. med. Wiss.* 1865, 881.

Spinal Vasomotor Centres. The great fall of blood pressure observed after section of the cervical cord is not permanent. After a short time (especially in the cat), the pressure begins to rise, and if the animal be kept alive, may attain a height equal to that found in normal animals.

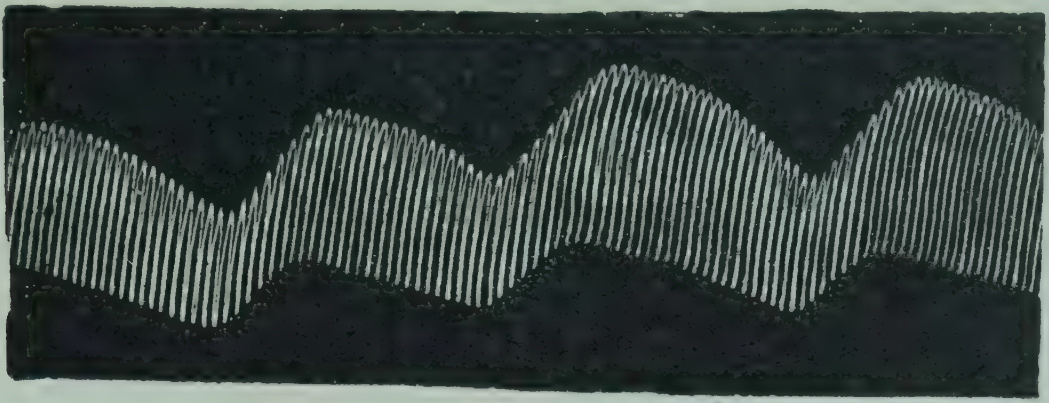


FIG. 451. Blood Pressure Tracing showing Mayer Waves. (C. J. MARTIN.)

If the spinal cord be now destroyed, the blood pressure sinks to a very low level, because the animal has been, so to speak, bled into its own dilated blood vessels. Subordinated to the chief vasomotor centre in the medulla, there is a series of subsidiary centres in the grey matter of the lateral horns of the cord, whence arise the fibres which enter the white rami communicantes. By means of these spinal centres a certain degree of adaptation of the blood supply to the various parts of the trunk is possible in the spinal animal. The important co-ordination between the state of the blood vessels and the rate of the heart is, however, wanting, since the spinal centres are now cut off from the medullary centres. In the intact body we must regard these spinal centres as subordinated to the higher vasoconstrictor centre in the medulla, forming, indeed, a relay on the outward path from these to the periphery; they are connected with the medullary vasomotor centre on each side by tracts of nerve fibres which descend in the lateral columns of the cord.

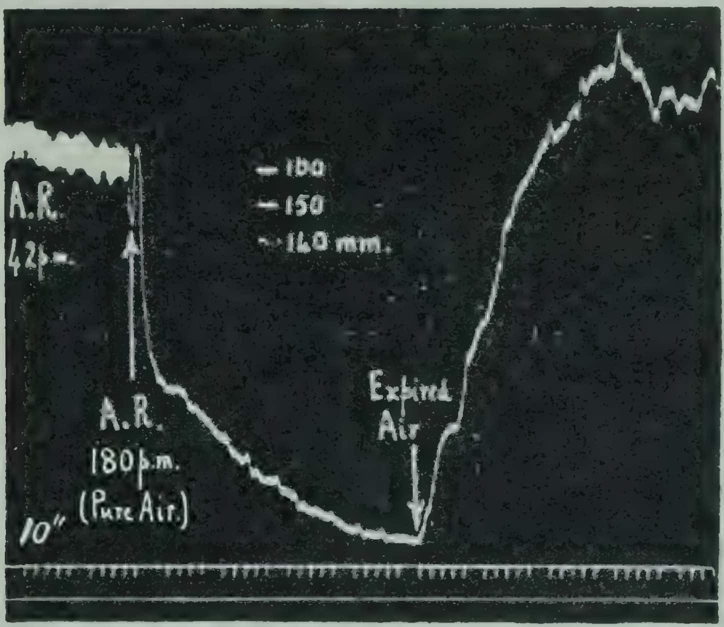


FIG. 452. Effect of Excessive Ventilation on Spinal Cat. Cord cut at second cervical level. With light artificial respiration (42 p. min.) the pressure is 160 mm.Hg; this rapidly falls on ventilation at 180 p. min., but is restored by respiration, at 180 p. min., with expired air. (DALE and LOVATT EVANS.)

The spinal centres are also susceptible to changes in the composition of the blood supplied to them. If an animal be kept alive by means of gentle artificial respiration after division of the cord just below the medulla, the blood pressure soon resumes a normal level. If artificial respiration be now discontinued the asphyxia excites the centres in the cord, and the pressure rises. Conversely, if the artificial ventilation be made excessive (acapnia), the blood pressure rapidly falls, but is speedily restored if ventilation, even at the same excessive rate, be continued with air containing 5 per cent. carbon dioxide, or with expired air (Fig. 452).* After destruction of the spinal cord, these effects disappear. The spinal centres are also excited by lack of oxygen. There is a difference between the sensitivity of the spinal and medullary centres; the common factor is probably increased H-ion concentration, but the excitation threshold for the medullary centres is lower than that of the spinal centres.

---

* AALKJÆR. *Skand. Arch. f. Physiol.*, 1935, **71**, 301.

THE PERIPHERAL TONE OF THE BLOOD VESSELS. Division of the sciatic and anterior crural nerves causes an immediate dilatation of the vessels of the lower limbs, in consequence of the loss of the tonic influence of the vasomotor centres. This dilatation diminishes in a day or two, and the vessels resume a tone, *i.e.* remain in a state of average constriction, which can be increased or diminished by local conditions. This recovery of tone has been ascribed to the tonus of the muscular fibres themselves. Division of the sympathetic supply to a limb (sympathectomy) is sometimes carried out as a surgical procedure. The blood flow in the limb is raised at first, but often returns to normal some months later; after ten years it is normal and the arterioles are hypertrophied.†

Vasomotor Reflexes. The vasomotor centre is constantly in receipt of impulses arriving both from the vascular system and from various other parts of the body. Hence reflex effects are produced by stimulation of various afferent nerves, and these may be classified, according to the area they affect, as general and local. If the reflex is sufficiently widespread, it causes a rise or fall of arterial pressure; such reflexes are distinguished as *pressor* or *depressor* reflexes respectively.

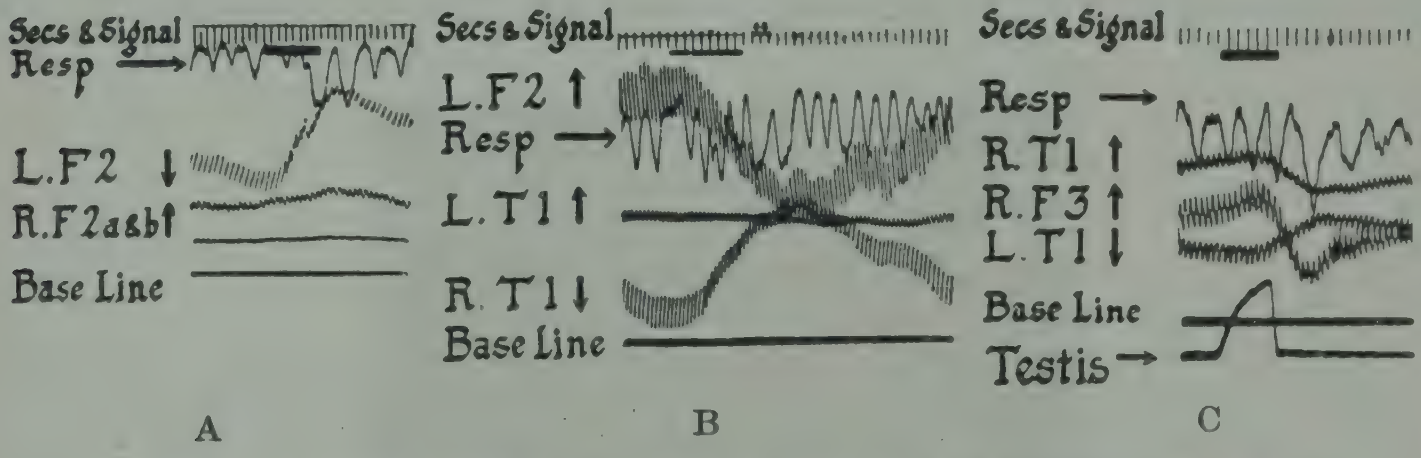


FIG. 453 Volume records of fingers or toes, and of respiration, in man.

R.F., right finger; L.F., left finger; R.T. and L.T., the right and left toes. The arrows show direction of dilatation.

A, lesion of right brachial plexus. Left arm pinched. No response of R.F.

B, left lumbar sympathectomy. Right arm pinched. No response in left toe.

C, lesion of cauda equina; motor and sensory paths to toes abolished, but sympathetic supply intact (from lower dorsal roots). Pressure on testis causes normal reflex. (STÜRUP *et al.*, 1935.)

Vasoconstrictor Reflexes. A *pressor reflex* is one which induces a rise of general blood pressure by widespread constriction of the arterioles, especially in the splanchnic area. Practically all impulses, which, if consciousness were present would be attended with pain or emotion, also cause vasoconstriction, which, if extensive and not otherwise compensated, causes a rise of arterial pressure. Thus a rise of pressure may be produced by the stimulation of almost any afferent nerve, including the central end of the splanchnic nerves.

These vasoconstrictor responses and the nervous pathways involved, have been elegantly demonstrated in man by Carmichael and his colleagues. One or more fingers or toes are placed in suitable plethysmographs and their volume curves are recorded. Stimulation of the skin, a loud noise, or the taking of a deep breath causes shrinkage of the digit by vasoconstriction.* During brief occlusion of the circulation in the arm by a sphygmomanometer armlet, the vasoconstrictor response still occurs in the fingers, showing that it is not a passive change. In cases of hemiplegia from lesions of the cerebral hemisphere the response in the affected side was normal.† Division of the whole nerve supply to the limb abolished the reflex, as did removal of the sympathetic supply alone (Fig. 453, A, B); division of the cerebrospinal supply without removal of the sympathetic supply left it unchanged (Fig. 453, C). Hence the path of the reflex in man is similar to that in the lower animals. It will be evident that the pressor reflexes are due to an increased sympathetic vasoconstrictor tone.

---

* DALE and LOVATT EVANS. *J. Physiol.*, 1922, **56**, 125.

† ESSEX, *et al.* *Amer. J. Physiol.*, 1943, **139**, 351.

So far as the general arterial pressure is concerned, the most important impulses arriving at the centre are those from the vascular system itself, especially from the heart, aorta and carotid sinus, and those from the higher parts of the brain. Whatever the condition of the heart, the brain always demands a high arterial pressure. A failing heart therefore indirectly evokes constriction of the systemic vessels, a fact which may lead to a vicious circle in cases where the heart is unable to empty itself against the resistance of the blood vessels.

An important part is played by these reflexes in enabling the body to counteract the effects of severe hæmorrhage. This, by causing a fall of general arterial pressure, would set in action the pressor aspect of both the aortic and carotid sinus reflexes, thus causing vasoconstriction and cardio-acceleration, so that the fall of pressure would be much less than it would be without this regulating mechanism.

Another aspect is that which concerns the effect of posture on the circulation (see later).

Vasodilator Reflexes. The chief examples of *depressor* nerves, viz. the depressor or aortic nerve itself, and the nerve of the carotid sinus, we have already studied in dealing with the heart reflexes. The carotid sinus apparatus is probably at least as important as the aortic depressor nerve, with which it normally collaborates. By their co-operation a rise of arterial pressure excites a depressor reflex, with cardiac slowing, thus reducing or checking the rise. The depressor reflex is mainly due to a suppression of vasoconstrictor tone, but

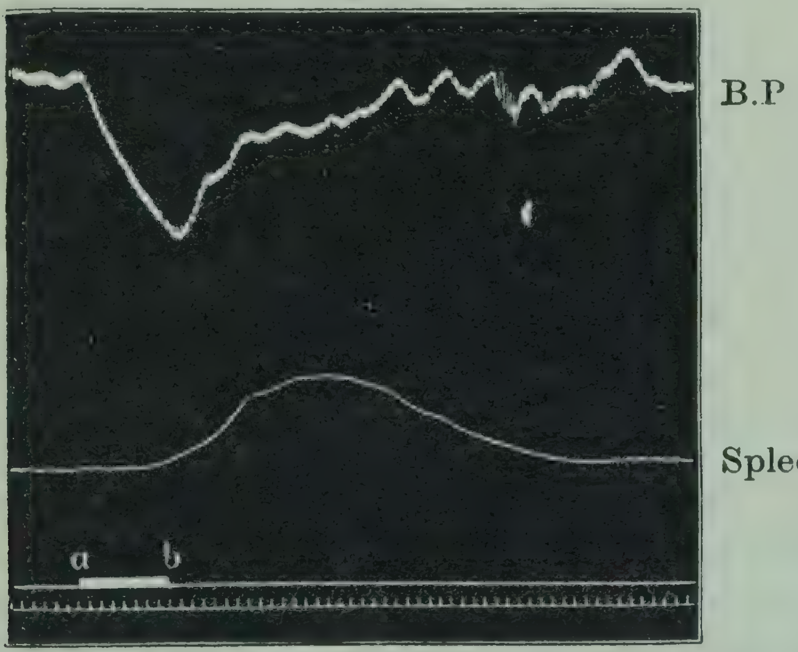


FIG. 454. Simultaneous Tracing of Arterial Blood Pressure and Splenic Volume from a Rabbit, showing the marked swelling of the spleen associated with fall of general blood pressure, on stimulation of the central end of the depressor nerve. The nerve was excited between *a* and *b*. (BAYLISS.)

---

* BOLTON, CARMICHAEL and STÜRUP. *J. Physiol.*, 1936, **86**, 83.

† STÜRUP *et al.* *Brain*, 1935, **58**, 456.

may be reinforced by the mediation of vasodilators, the two systems being reciprocally connected. The reflex fall of pressure produced by stimulation of the central end of the depressor nerve is effected chiefly by dilatation of the splanchnic area (Fig. 454), though, as Bayliss showed, practically all the vessels of the body partake in the relaxation.

Vasodilator reflexes arising from the carotid sinus and aortic zones may be reinforced by similar, but minor, reflexes arising from other vaso-sensitive zones, especially in the areas supplied by the cœliac and superior mesenteric arteries,* where the sensitive structures are the Paccinian corpuscles.†

Stimulation of the central end of the vagus, after section of both vagi, usually gives a depressor reflex because the nerve contains depressor fibres; but in some cases this may be preceded, or even replaced, by a rise of arterial pressure. This *vago-pressor* reflex appears to be normally in action when the blood pressure has been greatly lowered by hæmorrhage; in such conditions section of the vagi produces a further fall of arterial pressure (McDowall).‡

FAINTING. Often when the return of blood to the right heart is considerably reduced, fainting, or syncope, occurs. Thus it may happen after standing still for long periods, after considerable hæmorrhage, or when invalids attempt to stand. Fainting can be studied experimentally by performing a large venesection (1 l.), or by combining a smaller bleeding (500 c.c.) with "bloodless hæmorrhage." This is done by placing sphygmomanometer cuffs round the thighs of the sitting subject, and inflating to diastolic pressure. The legs become congested by retention of up to 700 c.c. of venous blood, which is thus removed from general circulation.§ After the faint has occurred, release of the cuffs restores this blood to circulation and the subject to consciousness. Prior to the attack there has usually been a steady rise in pulse-rate, and a definite decline of arterial blood pressure, and this, at or immediately before the faint, suddenly slumps to a low level, accompanied as a rule by a considerable drop in pulse-rate. The subject at the same time becomes pale, and perspires. The attack was called the "vasovagal syndrome" ‖ and was likened to the effects of stimulation of the carotid sinus, but its explanation is not so simple as that.

Closer study showed that as bleeding proceeded prior to the faint, the cardiac output and right auricular pressure steadily fell. The arterial pressure also declined somewhat, but calculation showed that the total peripheral resistance (= B.P./C.O.) increased considerably. This means that up to this point the falling arterial pressure caused by lowered cardiac output is opposed by widespread vasoconstriction and cardiac acceleration due to reduction of carotid sinus and aortic reflexes. But suddenly (Fig. 455), this opposition seems to melt away, the heart rate falls steeply, the peripheral resistance drops, and consciousness is lost because the blood supply to the brain becomes inadequate. The cardiac output may fall, or be unaltered, or even show some increase. The drop in heart rate is due to vagal inhibition: the fall of blood pressure is largely due to sudden extensive vasodilatation. Search for the site of the vasodilatation by plethysmographic methods (p. 627), showed that it is in the muscles, since the forearm blood-flow showed an increase (Fig. 456), but that of the hand (skin), a reduction.

Finally, in subjects with a sympathectomised arm, the forearm dilatation does not occur in that arm. It would appear, therefore, that sympathetic vasodilators (*v.* p. 646), convey the impulses to the muscle vessels.¶ The afferent path of the reflex is uncertain.

COMPENSATORY ADJUSTMENTS. Very commonly a vasoconstrictor reflex in one part of the body is accompanied by a vasodilator reflex in other areas, so that the arterial pressure may show no change, or may even fall. Under normal circumstances, every part of the body receives just enough blood for its metabolic requirements. Hence, activity must be associated with an increased flow of blood through the part. Two mechanisms are involved in this adaptation. In the first place, there is reflex dilatation of blood vessels in the part concerned, and in the second place, constriction of the blood vessels in the rest of the body, so that a normal or raised blood pressure is available for driving an increased supply of blood through the dilated vessels of the part. Thus, if both hind limbs of an animal be placed in a plethysmograph, it will be seen that stimulation of the central end of the anterior crural or peroneal nerve in the left leg causes dilation of this leg and constriction of the leg of the other side.* At rest, the organs of the chest and abdomen contain more than half of the total quantity of blood in the body, so that very little change in the capacity of these organs suffices to furnish the extra supply of blood needed by any part during a state of increased activity.

---

* HEYMANS, BOUCKAERT, FURBER and HSU. *Amer. J. Physiol.*, 1936, **117**, 619.

† GAMMON and BRONK. *Amer. J. Physiol.*, 1935, **114**, 77.

‡ MCDOWALL. *J. Physiol.*, 1924, **59**, 41.

§ EBERT and STEAD. *J. Clin. Inv.*, 1940, **19**, 561.

‖ LEWIS. *B.M.J.*, 1932, i, 873.

¶ BARCROFT and EDHOLM. *J. Physiol.*, 1945, **104**, 161.

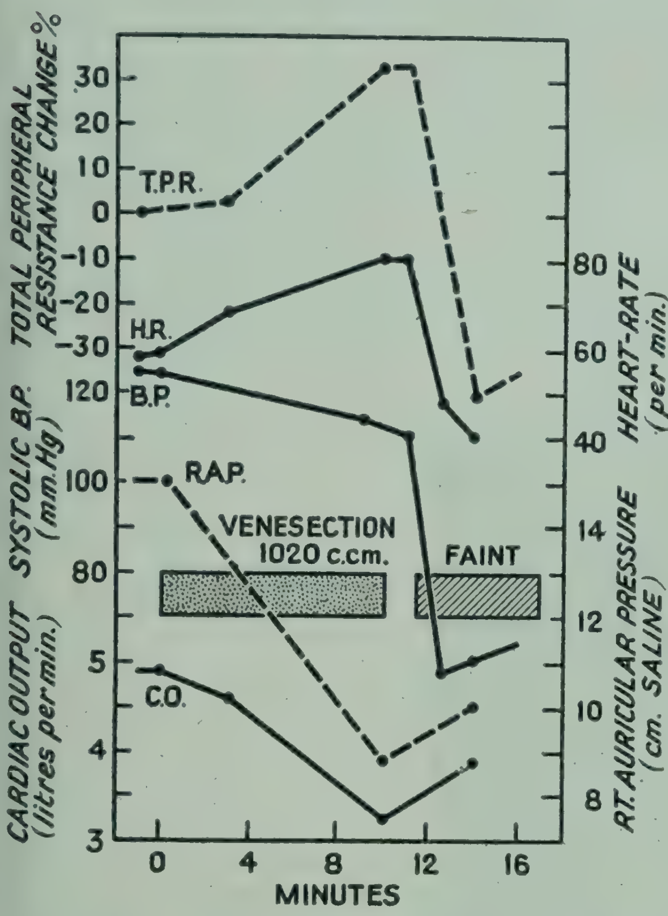


FIG. 455. Faint induced by venesection. CO., cardiac output; RAP., right auricular pressure; BP., systolic arterial pressure; HR., heart rate; TPR., calculated total peripheral resistance. (BARCROFT *et al.* *Lancet*, 1944, (i), 489.)

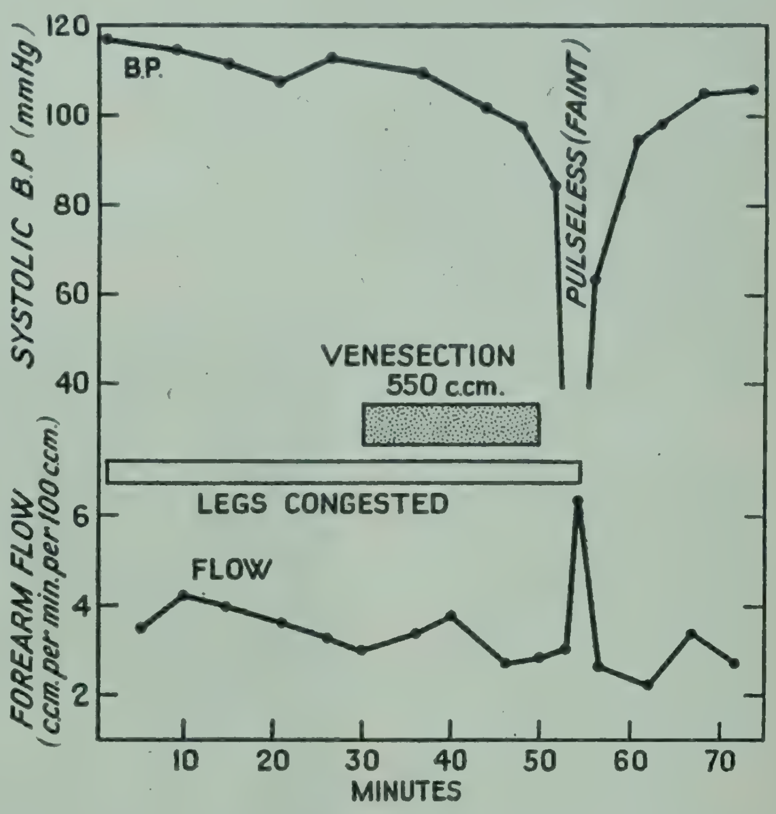


FIG. 456. Faint induced by venesection and leg congestion. Increase in muscle flow is considerable, and is promptly restored to initial level by release of cuffs round the legs. (BARCROFT *et al.* *Lancet*, 1944, (i), 489.)

The local axon reflexes, referred to above, also play an important part in bringing about adjustments of the blood flow in a part, and it has been shown by Fleisch that the reflex vasodilatation associated with activity of a tissue is in part a reflex due to chemical stimuli, as will be explained later. In the control of the blood supply to the tissues, the constriction and relaxation of the capillaries also probably play a large and special part.

These regulating reflexes provide yet another example, if any be needed, of one of the most fundamental of biological laws, namely that of the maintenance of constant conditions in the internal environment of the tissues, whereby a small change in the conditions gives rise to reactions which oppose further change, and oppose it with a force which increases as the change to be opposed augments.

---

* LOVÉN. *Ber. sächs. Ges.*, 1866, **18**, 85.

Cephalic Pressure. The head regulates its blood supply according to its needs. If the head be supplied with blood by means of a circulation distinct from that of the rest of the body, it is found that any rise of pressure in the arteries of the head brings about the inverse change in the arteries of the rest of the body—so that rise of pressure in the head vessels means a fall of pressure in other parts of the body, and *vice versâ*.* This phenomenon, like the reflex cardio-inhibition which accompanies it, is a reflex from the *carotid sinus* (*v.* p. 595),† and is not a consequence of the effect of pressure on the brain itself. The effect is reduced or abolished by denervation of the carotid sinuses.

For Reference

Bayliss (1923). The Vaso-motor System. (Longmans, Green.)
Burn (1938). Sympathetic Vasodilator Fibres. *Physiol. Rev.*, **18,** 137.
Heymans, Bouckaert and Regniers (1933). Le sinus carotidien, etc. (Paris: Doin.)
Lewis (1927). The Blood Vessels of the Human Skin and their Responses. (London: Shaw & Son.)
McDowall (1935). The Nervous Control of the Blood Vessels. *Physiol. Rev.*, **15,** 98.

THE CHEMICAL REGULATION OF THE ARTERIOLES

Vasodilator Substances. Metabolites. A factor involved in the production of the increased blood flow through active organs is a chemical stimulation of the vessels themselves, by means of substances (metabolites) produced as a

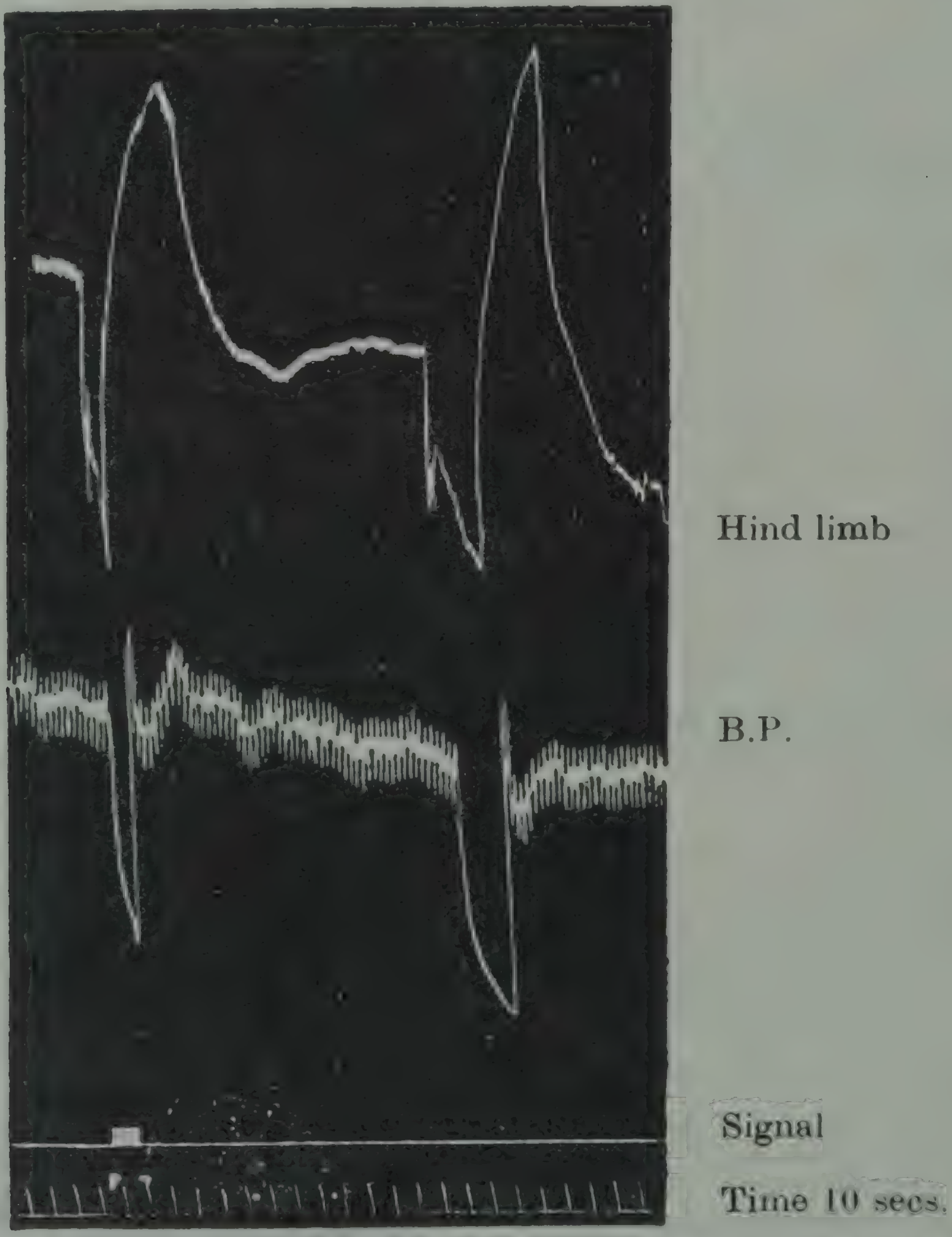


Fig. 457. Effect of Temporary Compression of the Abdominal Aorta on the Volume of the Denervated Hind Limb.

Two compressions, the second not marked by the signal. Blood pressure taken in the femoral artery of one hind limb, the other hind limb being in the plethysmograph. (Bayliss.)

---

* Anrep and Starling. *Proc. Roy. Soc.*, 1924, **97B,** 463.

† Florey, Marvin and Drury. *J. Physiol.*, 1928, **75,** 209. Heymans. *Le sinus carotidien*, 1929.

result of activity. These substances are continually being produced, though in smaller amount, by the resting tissues, and when the blood supply is stopped for a time, they accumulate. If the blood is readmitted to the tissue, its vessels widely dilate—a condition called *reactive hyperœmia*. This is illustrated by Fig. 457. The great increase in the flow through the muscles, which accompanies muscular exercise, is probably brought about largely by similar means.

The exact nature of the metabolites concerned is still uncertain. Lactic acid and carbon dioxide cause some dilatation of the blood vessels of a limb, but the dilatation in activity is greater than could be accounted for by these substances. Several other possible metabolites have been found to exert a dilator effect on the arterioles. Thus, it may be (1) acetyl choline, or some closely related substance; when a vasodilator nerve or an antidromic sensory nerve is stimulated, there is a liberation, in the region of its distribution, of acetyl choline; (2) adrenaline has a vasodilator action on some vessels, *e.g.* in man a transient one on those of muscles; (3) adenosine, adenylic acid or adenyl pyrophosphate are all common in tissues, but particularly in muscle; (4) some not yet identified substance, such as Gaddum and Euler's * "substance P," extractable from intestine and brain, or "kallikrein," present in urine, and various extracts of protein nature.†

It might be asked, however, why the arterioles should be affected by a substance which enters the blood down-stream from them, in the capillaries. The explanation, according to Fleisch,‡ is that an ascending local reflex is effected through the nerves in the vessel walls. These nerves may be supposed to be stimulated by the various metabolites, and by a local reflex to cause arterial dilatation.

Vasoconstrictor Substances. (1) ADRENALINE. This substance has a powerful vasoconstrictor action: if 1 c.c. of a 1 in 10,000 solution be injected into the jugular vein, there is within a minute a widespread constriction of the arterioles, with the exception of those of the muscles, heart and brain. The same results occur after the nervous system has been destroyed, proving that adrenaline acts directly on the blood-vessels. Owing to the large rise in arterial pressure, the action is therefore to divert the supply of blood from skin and internal viscera, and to augment the flow through muscle, brain and heart.

When the splanchnic nerve is stimulated, there is a release of adrenaline, and the rise of arterial pressure shows a stepped curve, the first rise being due to the action of the vasoconstrictor nerves on the blood vessels, the second to the discharge of adrenaline into the general circulation (Fig. 458). Similarly in Fig. 459 is shown the effect of brief stimulation of the splanchnic nerve on the blood pressure and on the volume of the limb of the cat. It will be noticed that the volume of the hind limb first increases passively with the rise of pressure, and then diminishes to much below its previous volume. This diminution is due to the discharge of adrenaline into the blood stream and is absent if the suprarenals have been removed. During asphyxia, the rise of arterial pressure is associated with the discharge of adrenaline. Similarly, when any pressor reflex is set up, as, *e.g.* by a reduction of tension in the carotid sinus, it is accompanied by an increased discharge of adrenaline into the circulation.

The most intense action of adrenaline is on the skin vessels. The arterioles of muscle are transitorily dilated, and those in the small intestine only feebly constricted by adrenaline, while the coronary vessels of the heart are intensely dilated by it.

(2) NOR-ADRENALINE (*v.* p. 361), has still more powerful vasoconstrictor actions and no vasodilator actions. It is formed and liberated together with adrenaline, and probably in about equal amount.

(3) RENIN AND HYPERTENSIN. In 1898, Tigerstedt and Bergman observed

---

* EULER and GADDUM. *J. Physiol.*, 1931, **72,** 74.

† FREY and KRAUT. *Ztschr. f. Physiol. Chem.*, 1926, **157,** 32.

‡ FLEISCH. *Arch. Internat. de Physiol.*, 1935, **41,** 141.

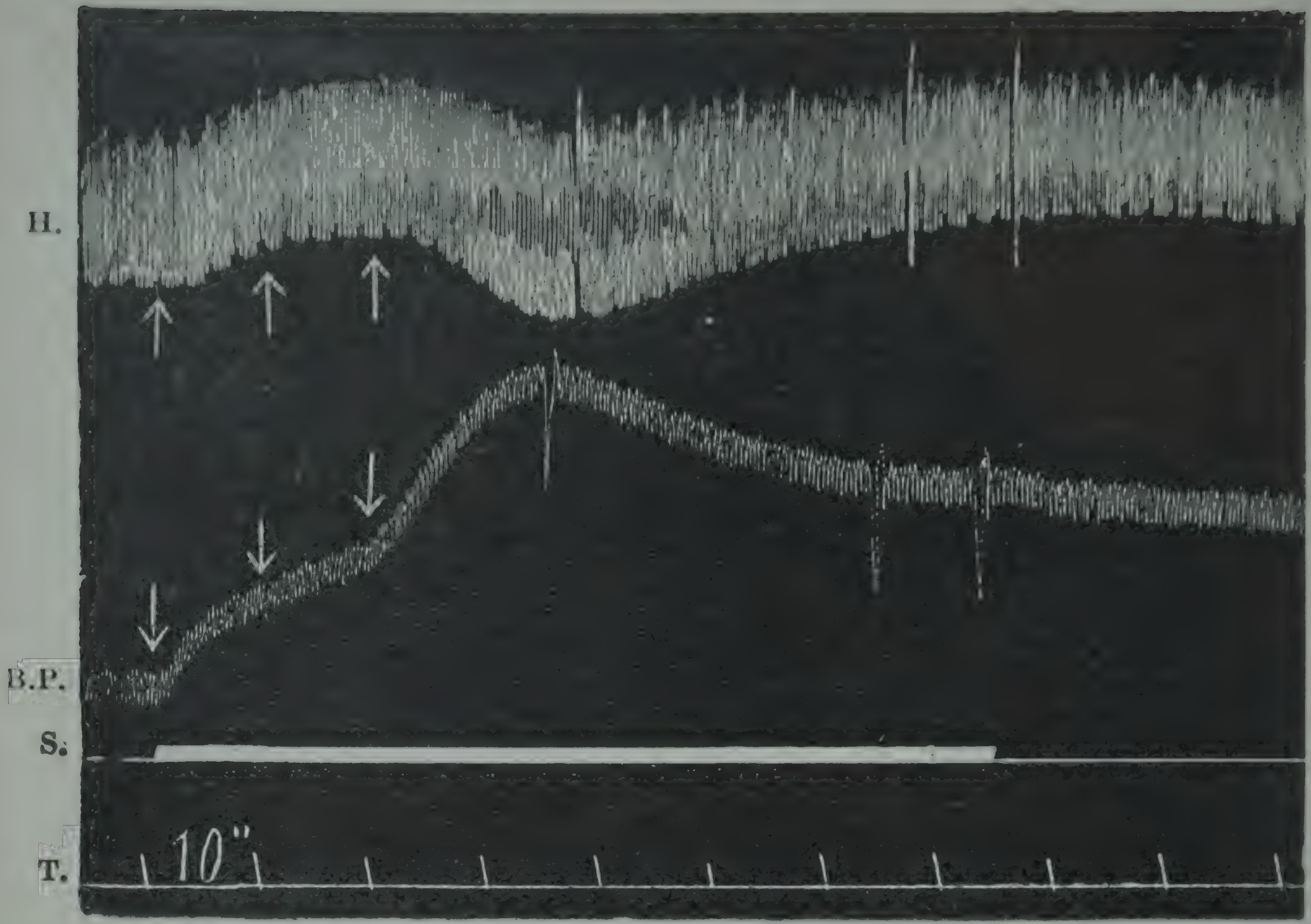


Fig. 458. Effect of Splanchnic Stimulation on Arterial Pressure and on Heart Volume.

H. Heart volume. (Systole causes a downward movement of the lever.) B.P. Arterial blood pressure. S. Stimulation of splanchnic nerve. T. Time, 10 secs.

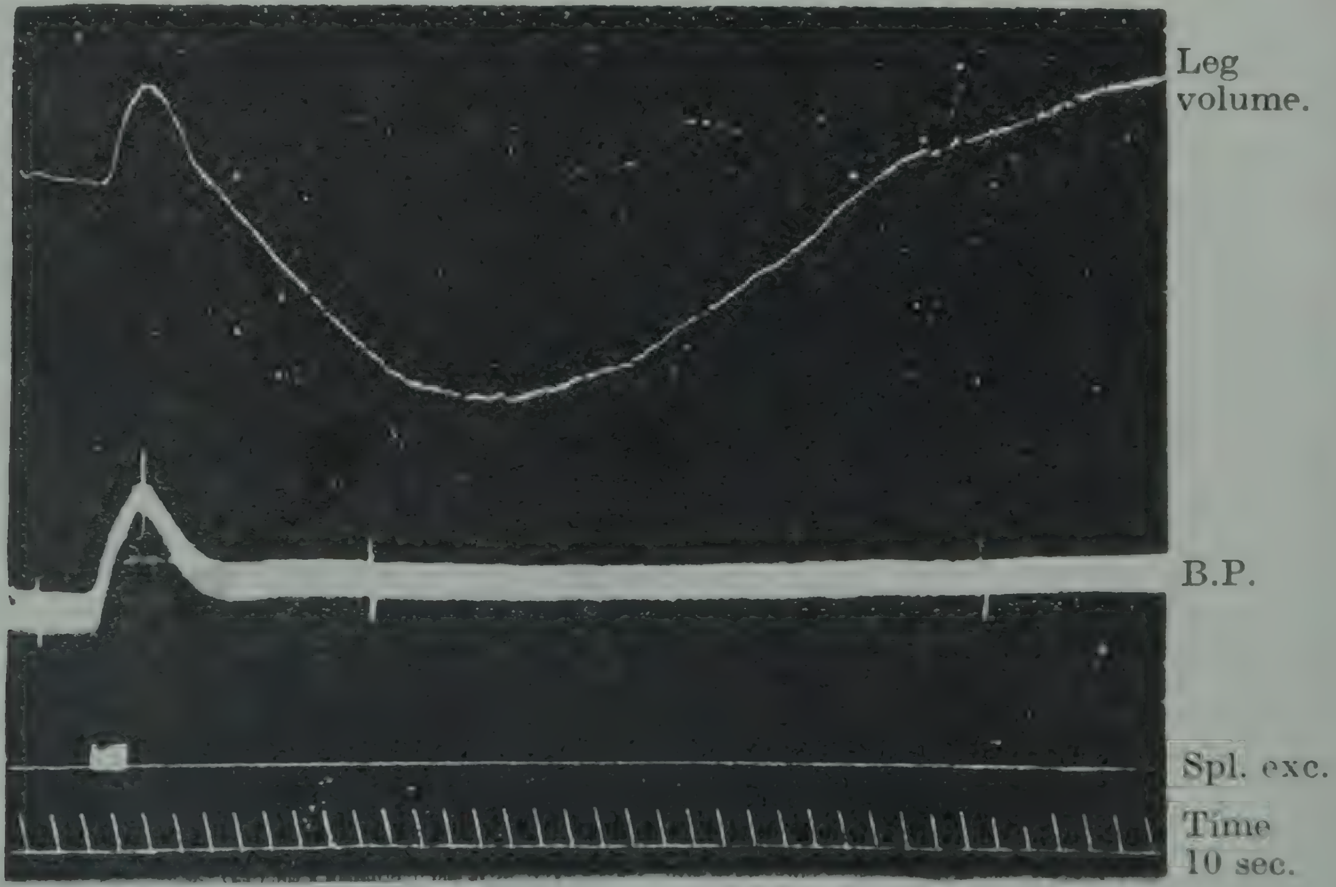


Fig. 459. Effect of Excitation of Splanchnic Nerves on the Blood Pressure and on the Volume of the Denervated Hind Limb of the Cat. (Bayliss.)

a prolonged rise of arterial pressure after intravenous injection of extracts of kidney cortex and gave the name *renin* to the responsible substance.* An interesting story has eventuated, which provides a possible hypothesis for the ætiology of the hypertension met with in chronic nephritis. Goldblatt [*] showed that if the renal blood flow is restricted by constriction of the renal artery, a pressor substance is released from the kidney into the circulation, and a persistent hypertension results. The extractable substance, renin, is not, however, itself a pressor substance : it is an enzyme, which acts on a plasma globulin ("*hypertensinogen*") to produce a pressor agent called "*hypertensin*" [†] which appears to be a polypeptide or proteose.[‡] The vasoconstriction produced by hypertensin is widespread, but affects the skin relatively little. The renal blood flow is reduced ; the substance is slowly destroyed in the body by an enzyme. Its possible function in the body is to prevent urine formation from failing when the reduced access of blood to the kidney threatens this : it constricts the efferent arterioles of the glomeruli, so keeping up the intraglomerular pressure.

---

* Pronounced "*reenin*," for obvious reasons.

(4) VASOTONINS. These are constrictor substances, of unknown nature, produced in blood when it clots, and perhaps continually in the body at a slow rate. They are therefore present in defibrinated blood, from which they can be removed by perfusion of the blood through the vessels of excised lungs.

THE REGULATION OF THE BLOOD FLOW THROUGH THE CAPILLARIES

Although the arterial blood pressure and the state of contraction of the arterioles play a large part in regulating the supply of blood to the tissues, we cannot regard the capillaries as merely inert channels. The calibre of the capillaries is determined to a great extent by changes in their own walls ; these changes consist in a swelling or shrinkage of the nuclei of the endothelial cells whereby the lumen of the capillary is narrowed or widened.[§]

When a tissue is in a state of rest, many of its capillaries close down completely, and so reduce its blood supply to a level commensurate with its small oxygen requirement. On activity, as the oxygen needs of the tissue increase, more and more capillaries open up, so that in a tissue like muscle there may be at least a hundred times as many capillaries in use as there were in the resting state.

Capillaries can be caused to contract by direct stimuli, or through the action of nerves, or by alteration in their chemical surroundings.

Action of Nerves. The capillaries are abundantly supplied with nerve fibres, and capillary constriction has been observed under the microscope on stimulation of the sympathetic nerve supply.[||] There is also a connection between nerve lesions and the calibre of the capillaries in the area supplied by the nerves. Thus, if in the cat the sciatic nerve be cut on the right side, for the next few hours the pad of the foot on that side is flushed and warmer than the left foot. The next day the flush has disappeared, in fact the pad of the right foot may be paler, though still warmer, than that of the left. This may be because, although the arterioles are still dilated and arteriovenous anastomoses open, so that there is a more rapid flow of blood through the pad, the capillaries are contracted.

It would be expected that section of the sympathetic fibres, by causing dilatation, would make the capillary walls more permeable. But actually the reverse appears to be true, *i.e.* the dilated sympathectomised capillaries are less permeable than normally.[¶]

Whether there are dilator nerves to the capillaries is uncertain, but some claim that axon reflexes are demonstrable, similar to those affecting arterioles.

---

[*] GOLDBLATT. *Harvey Lects.*, 1938, **33**, 237.

[†] BRAUN-MENENDEZ *et al.* *J. Physiol.*, 1940, **98**, 283.

[‡] EDMAN *et al.* *J. Physiol.*, 1942, **101**, 384.

[§] SANDERS, EBERT and FLOREY. *Q. J. Exp. Physiol.*, 1940, **30**, 281.

[||] BEECHER. *Skand. Arch. f. Physiol.*, 1936, **73**, 123.

[¶] ENGEL. *J. Physiol.*, 1941, **99**, 161.

Skin Reactions. The independent contractility of the minute vessels is well illustrated in the human skin, the colour of which is determined chiefly by the number and size of its minute vessels ("capillaries" or "sub-papillary venous plexuses").

If a blunt point be drawn over the skin, a mark, or *tache*, is left behind. On gentle stroking or on applying pressure with the finger, the *tache* is white, and appears in fifteen to twenty seconds, being preceded by transient flushing. When the point is drawn over the skin with a firmer pressure, there is first a light flush (capillary dilatation), followed in fifteen seconds or so by (*a*) a much deeper flush along the actual track of the instrument (wide dilatation of capillaries and other vessels), and (*b*) at either side of the red *tache* so produced there appears a white *tache* (capillary constriction). These fade in from three to eight minutes. Microscopic inspection shows that the capillaries and venules in the pale area, though constricted, are in connection with vessels filled with blood, and the emptiness of these vessels can only be due to the fact that they have actively contracted against the pressure of blood with which they are supplied.

An experiment by Lewis, called the occlusion test, conclusively demonstrates capillary contractility. A Riva-Rocci armlet is placed in position and *rapidly* inflated to a pressure greater than the arterial pressure, *e.g.* 200 mm. Hg. The circulation in the arm is thus suddenly stopped, and the pressure in arteries, capillaries and veins then becomes equal; as the inflation was rapid, this pressure, though higher than the usual capillary pressure, is not very great. On now applying the skin reaction test, the white *tache* will be found to develop in the usual manner. Now, since the colour of the skin is due to blood contained in its minute vessels, it is evident that in such a limb, with occluded vessels, a contraction of arterioles could not deprive the capillaries of their contained blood, as might be objected if the circulation were proceeding.

Lewis has shown that the red reaction consists of two distinct reactions—a localised one, due to relaxation of *all* the skin vessels, and a spreading flare, due to an axon-reflex, which causes relaxation of the skin arterioles, often for considerable distances, when powerful stimuli have been used.

In a certain number of subjects, the skin after the *tache* tests shows some urticaria, while in others a definite wheal results, owing to the production of local œdema—a condition called *dermatographism.* The *triple response, i.e.* local vasodilatation, spreading flare and local œdema, may be obtained by stimulation of the skin by mechanical means, heat, cold, electricity and by various chemical substances, especially histamine.

Whether these reactions characteristic of the minute vessels of the skin are properties of the capillaries in other regions is uncertain. It would be expected that the skin would be best adapted to show responses to mechanical stimulation, but there is a certain amount of evidence for similar responses in other parts to stimuli equally appropriate.

METABOLITES. The dilator effect of metabolites on the arterioles, which we have already studied, seems to be shared by the capillaries. In such a case, it is difficult to dissociate the effects of arterial dilatation from those of capillary dilatation. Acetyl choline, or substances akin to it, chiefly relaxes the arterioles. At least one chemical substance is known, however, which has diametrically opposite effects on the two sets of vessels, viz. *histamine.* This substance has been shown to have a slight constrictor effect on the arterioles and a powerful dilator effect on the capillaries.* Its effects are antagonised by the drug *benadryl.* Moreover, it so increases the permeability of the capillary walls, that fluid rapidly passes out from the blood stream into the tissues. It has been suggested that when tissues are subjected to damage, histamine or something resembling it (H-substance) is either liberated from or produced in the cells, and results in local capillary and venule dilatation as seen, for instance, in the skin reactions. If a minute amount of histamine is introduced beneath the skin by pricking through a drop of 1 : 1,000 solution, a typical triple response is seen at the site.

---

* DALE and RICHARDS. *J. Physiol.*, 1918, **52**, 110.

CIRCULATORY SHOCK

Circulatory shock is a state of circulatory failure, which results from loss of blood, or extensive damage or prolonged suspension of blood supply to the tissues, particularly muscle. It is seen after severe surgical handling, bomb damage, burns, prolonged application of tourniquets to limbs, etc., and is studied in the laboratory by applying tourniquets for several hours to the limbs of anæsthetised animals, and then removing them, or by inflicting crush damage to the limbs or trunk. All the effects are magnified if hæmorrhage also occurs.

The essential features are lowered venous pressures, and diminished venous return, so that the cardiac output is reduced. The prime factor in its causation appears to be loss of circulating fluid, whether of blood by external or internal hæmorrhage, or of plasma as in burns, or in " toxic " shock ; or of blood plus plasma in crush injuries. Ultimately a vicious circle is set up in the peripheral circulation, with deficient tissue oxygen supply, widespread capillary dilatation, and increased capillary permeability, resulting in a continually increasing reduction in blood volume, due to loss of plasma. Consequential phenomena are, " hæmoconcentration," increase of blood viscosity ; vasoconstriction of arterioles, which reduces the supply of blood to the tissues, but which may keep the arterial pressure from falling until the loss of blood volume reaches some 25 per cent. ; fall of venous pressure, which reduces cardiac output ; oxygen lack of the tissues leading to a fall of temperature, production of lactic acid, fall in alkali reserve, etc., so completing the vicious circle. The heart is accelerated as a compensatory effect, but ultimately the arterial pressure falls so low that the heart weakens and finally fails from oxygen lack. There may be kidney damage, due to low arterial pressure. [Another kind of kidney damage may be present after extensive crushing injuries : it may be due to the setting free of myoglobin from injured muscle, and this, when acid urine is excreted, leads to damage of the kidney tubules and renal failure.*]

Three types of theory have sought to explain this complicated syndrome. (1) That part, at least, is of reflex origin, and leads to exhaustion of the vasoconstrictor centres, or of the suprarenal medulla. It has been stated that if the nerves from the damaged tissues are divided, or blocked by novocaine, shock is reduced or absent. Most of the force of this argument is destroyed by the clear proof that the vasoconstrictor apparatus and the adrenal medulla function quite well, even at advanced stages of shock. (2) That loss of fluid by its retention in the damaged area may be enough to cause the results. But if a limb is clamped off for seven hours, and the clamp released, amputation within the next two hours may prevent death, though even after further time the damaged limb shows no great increase in weight. (3) That the source of the trouble is chemical. Thus, lapse of function of the suprarenal cortex has been suspected, and administration of corticosterone advised. The chief chemical view is that some toxic substance arises in the damaged tissues and raises the permeability of the capillaries throughout the whole body.

Search for the responsible chemical agent has only met with partial success. That the effects may be due to the liberation of histamine, or some closely-allied substance, is mainly circumstantial. For instance, histamine can, on injection into the circulation, often cause shock ; or, when injected into the tissues, cause local changes similar to those produced by local injury. There are reasons, however, for not accepting all this evidence at its face value.†

---

* Bywaters and Stead. *Q. J. Exp. Physiol.*, 1944, **33,** 53.

† Grant and Duckett. *Heart*, 1929, **14,** 337 ; Rous and Gilding. *J. Exp. Med.*, 1930, **51,** 27 ; Percival and Scott. *J. Pharm.*, 1931, **41,** 147.

There is no evidence that the histamine content of the blood from a traumatised part, or of that in general circulation in shock patients, is much, if at all, increased

A more promising approach attributes the effects to the liberation, especially from muscle, of a substance, or substances, called the muscle shock factor(s).*

Again, there is some evidence that Kallikrein may be the responsible substance.†

For Reference (Capillaries, Vasomotor, and Chemical Control)

BEST and MCHENRY (1931). Histamine. *Physiol. Rev.*, **11,** 371.

BLALOCK (1940). Experimental Hypertension. *Physiol. Rev.*, **20,** 159.

BRAUN-MENENDEZ *et al.* (1943). Hypertension arterial nefrogena. (Buenos Aires: "El Athenes.")

DRURY (1936). The Physiological Activity of Nucleic Acid and its Derivatives. *Physiol. Rev.*, **16,** 292.

GADDUM (1936). Gefässerweiternde Stoffe der Gewebe. (Leipzig, Thieme.)

GOLDBLATT (1947). The Renal Origin of Hypertension. *Physiol. Rev.*, **27,** 120.

HEYMANS, BOUCKAERT and REGNIERS (1933). Le sinus carotidien et la zone homologue cardio-aortique. (Doin: Paris.)

HOSKINS (1922). The Relation of the Adrenals to the Circulation. *Physiol. Rev.*, **2,** 343.

KROGH (1929). The Anatomy and Physiology of Capillaries. (Yale Univ. Press.)

LANDIS (1934). Capillary Pressure and Capillary Permeability. *Physiol. Rev.*, **14,** 404.

LEWIS (1927). The Blood Vessels of the Human Skin and their Responses. (London: Shaw and Son.)

MCDOWALL (1935). The Nervous Control of the Blood Vessels. *Physiol. Rev.*, **15,** 98.

MCDOWALL (1938). The Control of the Circulation of the Blood. (London: Longmans.)

MOON (1938). Shock and Related Capillary Phenomena. (London. Oxf. Univ. Press.)

SCHMIDT and COMROE (1940). Functions of the Aortic and Carotid Bodies. *Physiol. Rev.*, **20,** 115.

WIGGERS (1942). The Present Status of the Shock Problem. *Physiol. Rev.*, **22,** 74.

* GREEN. *Lancet*, 1943, **245,** 147; BIELSCHOWSKY and GREEN. *Lancet*, 1943, **245,** 153.

† WESTERFELD *et al.* *Amer. J. Physiol.*, 1944, **142,** 519.

CHAPTER XXXV

THE PULMONARY, HEPATIC AND CEREBRAL CIRCULATIONS

THE PULMONARY CIRCULATION

In the lungs the alveoli are supplied by a close network of wide arterioles and capillaries carrying venous blood derived from the pulmonary artery, and presenting very little resistance to the flow of blood ; there is a much smaller supply of arterial blood, to the bronchial tubes, by way of the bronchial arteries. There is a certain amount of anastomosis between the pulmonary and bronchial circulations, though mainly only in the smaller vessels, so that probably but little bronchial blood reaches the alveolar capillaries.* The mean pressure in the pulmonary artery probably does not exceed about one-sixth of the mean aortic pressure. By catheterizing branches of the pulmonary artery and vein in dogs with thorax intact, the pressures were found to be 10 mm. Hg and 6 mm. Hg respectively : the pulmonary capillary pressure is therefore about 8 mm. Hg.†

The capillaries of the lungs may vary passively in size, according to the conditions. Thus, whereas at the height of inspiration the blood contained in the lungs is about one-twelfth of the whole blood in the body, this amount is said to be diminished during expiration to between one-fifteenth and one-eighteenth, and by forcible artificial inflation of the lungs may be lessened to one-sixtieth. Such changes may exercise a considerable effect on the systemic blood pressure. Again, if the circulation rate is increased, the whole of the pulmonary vascular bed will be distended, and as much as one-fifth of the blood may be in the lungs (Kuno ‡). On the other hand, the distensibility of the lung capillaries may play an important part in enabling the lungs to act, so to speak, as a reservoir for the left side of the heart. If there is a temporary excess of output on the right side that cannot be dealt with at once by the left heart, the excess is taken up for a time in the lung vessels.

The mean pressure in the pulmonary artery depends on the output of the right ventricle and the resistance in the lungs, and of these the one most subject to variation is the former. The mean pressure in the pulmonary veins is very low, being at or about zero, and somewhat greater than that in the left auricle. Because resistance is low, the gradient between pulmonary artery and pulmonary vein is not so steep as might be expected in view of the high velocity of the blood in the pulmonary circuit. Although an increased output of the right heart leads to a rise of pressure in the pulmonary artery, the pressure in the pulmonary veins is little increased. This fact, which is because the left auricle is normally competent to keep the pulmonary veins adequately drained, results in steepening of the gradient of pressure which drives blood through the lungs.

An important question is raised as to the influence of changes of aortic pressure, due to alteration of systemic resistance, on the pressures in the pulmonary circuit. Experiments on the heart-lung preparation showed that a rise in the aortic pressure, while the inflow into the heart from the venous reservoir was kept constant, was accompanied by a rise in the pulmonary artery pressure. This was at first ascribed to " back pressure," *i.e.* to a relative failure of the left auricle to deal with the blood flowing into it. This cannot be the case, however, for the mean left auricular pressure is not raised (Fig. 460). The phenomenon is due to increased output of the right heart caused by increased flow through the coronary arteries,[*] and to a lesser extent through the bronchial arteries[†] induced by the rise of aortic pressure. In the closed-circuit heart-lung preparation, with constant ventricular output a rise of aortic pressure has little or no effect on pulmonary arterial pressure.[†] Similar conditions contribute to the parallelism between pulmonary and systemic arterial pressures in the intact animal. Only when the left ventricle is greatly dilated (and the mitral valve incompetent) are there likely to be pronounced back pressure effects.

---

* Berry and Daly. *Proc. Roy. Soc.*, 1931, **109B,** 319.

† Hellems *et al.* *Amer. J. Physiol.*, 1948, **155,** 98.

‡ Kuno. *J. Physiol.*, 1917, **51,** 154.

The pulmonary arterial pressure shows considerable variations, about 30 mm. Hg with each heart beat, so that the pulse pressure is greater than the average pressure[‡]; this means that the resistance in the pulmonary bed is low. The pulsations resemble those of the systemic arteries (Wiggers). There are also slower variations of pressure due to the respiratory movements; the pressure falls during inspiration and rises during expiration. In the pulmonary veins the pressures vary less with respiration, and probably in the reverse direction. These results are due to a lowering of resistance in the pulmonary capillaries during inspiration.[||]

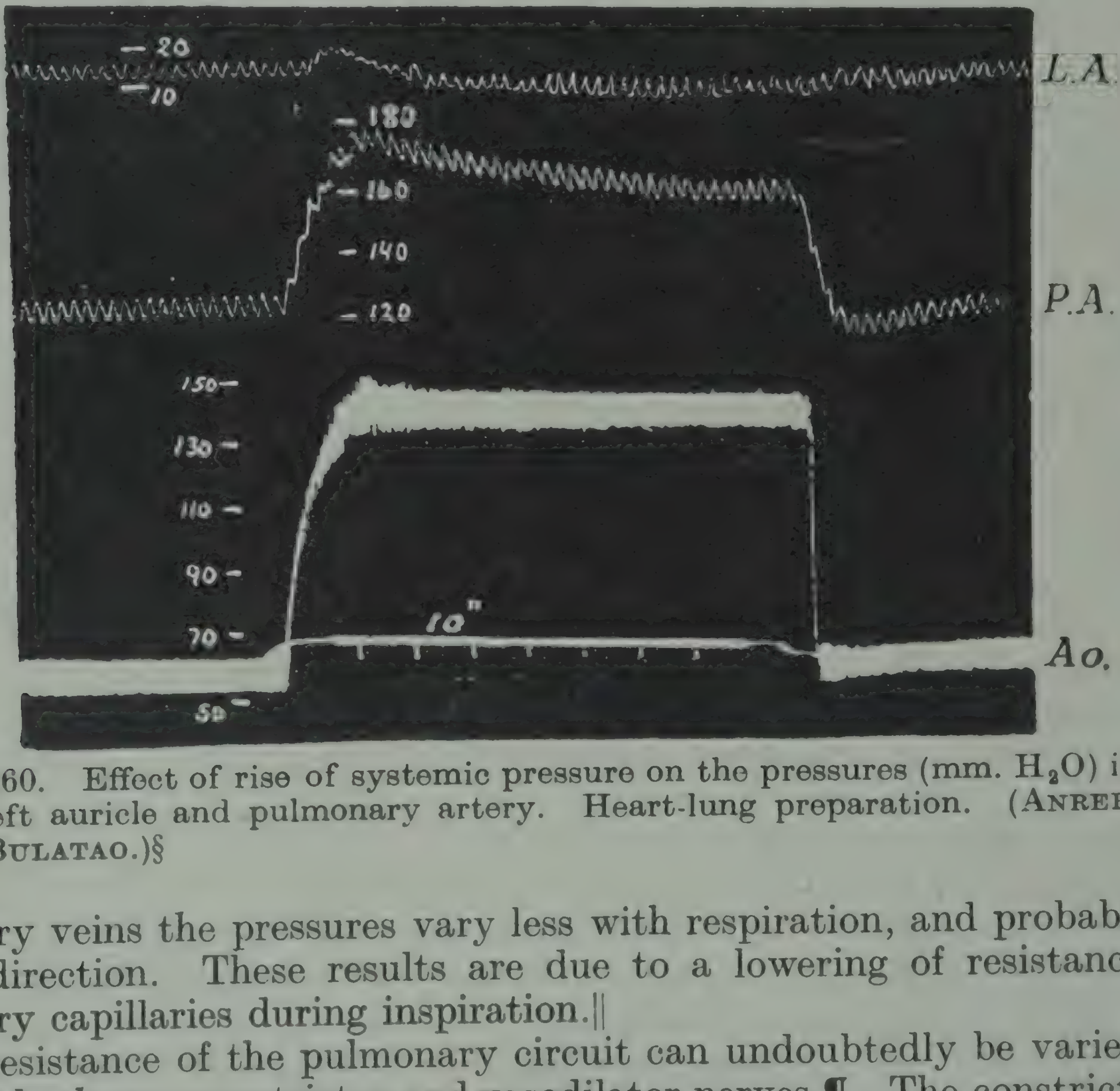


FIG 460. Effect of rise of systemic pressure on the pressures (mm. H_2O) in the left auricle and pulmonary artery. Heart-lung preparation. (ANREP and BULATAO.)[§]

The resistance of the pulmonary circuit can undoubtedly be varied by the action of both vasoconstrictor and vasodilator nerves.[¶] The constrictor fibres are sympathetic from the upper thoracic nerves, *via* the vago-sympathetic nerves. Addition of adrenaline to the circulating blood causes pulmonary vasoconstriction with rise of pulmonary artery pressure. The dilator fibres are probably derived from the vagi. Stimulation of the vago-sympathetic may therefore cause either vasoconstriction or vasodilatation, according to which of the component fibres are most affected. Reflex pulmonary vasoconstriction has been claimed to result as a consequence of the stimulation of sensory nerve roots and of the central end of the splanchnics, and reflex dilatation from stimulation of the depressor nerve. It is uncertain to what extent the nervous control of

[*] ANREP and BULATAO. *J. Physiol.*, 1925, **60**, 175.
[†] DALY. *Proc. R. S.* (B), 1926, **99**, 306.
[‡] MATTHES and HOCHREIN. *Arch. f. Exp. Path.*, 1932, **167**, 678.
[§] ANREP and BULATAO. *J. Physiol.*, 1925, **60**, 175.
[||] DALY. *J. Physiol.*, 1930, **69**, 238.
[¶] DALY and v. EULER. *Proc. Roy. Soc.*, 1932, **110**, 92; DALY and HEBB. *Q. J. Exp Physiol.*, 1942, **31**, 211.

the pulmonary bed is exercised under normal conditions or what its function may be, apart from altering the amount of blood in the pulmonary circuit.

Effect of Respiration on Systemic Blood Pressure. A tracing of the general arterial pressure shows that with each inspiration the blood-pressure rises, and with each expiration it falls. These undulations are in part due to an irradiation from the inspiratory centre to the vasomotor centre (p. 652) (Traube-Hering waves): in part therefore they are independent of mechanical changes in the thorax. The synchronism of the rise and fall with the respiratory movements is not exact, and varies with the rate of the respiratory movements, since the rise continues for a short time after the beginning of expiration, and the fall continues into the beginning of the next inspiration, so that in slow breathing the highest point of the curve occurs at the beginning of expiration and the lowest point at the beginning of inspiration. During the rise at inspiration, the heart beats often become accelerated. This is due to a "Bainbridge reflex," the acceleration being caused by the increased filling of the right heart by aspiration of blood into it during inspiration. On dividing both vagi, this difference in the pulse rate during inspiration and expiration disappears, but the main features of the blood pressure curve remain; when respiration is artificially produced by forcing air into the lungs, the effect of the movements is reversed, *i.e.* the rise now occurs during the expiratory phase; these various facts indicate that some mechanical factor must also be concerned in the explanation of the respiratory undulations.

As we shall see later, under normal conditions the lungs are in a state of stretch, and in consequence of this condition they are constantly exerting a pull on the chest wall. As the chest expands in inspiration it drags the lungs still more open, so that their pull on the chest wall becomes greater, and hence the negative pressure in the pleural cavity is much increased during forcible inspiration. The heart, great veins and arteries in the thorax are practically exposed to any pressure which may exist in the pleural cavity, whereas outside the thorax all the vessels are exposed to a positive pressure.

Any fluid will always flow from a point of higher to a point of lower pressure. Hence, when the thorax enlarges during inspiration, there must be an aspiration of blood into it from peripheral parts. This aspiratory force will not influence arteries and veins alike. The arteries, having thick walls, will be very little affected by the increase of negative pressure in the thoracic cavity, whereas the thin-walled distensible veins, auricles and right ventricle, will be largely influenced. The total result, then, of the increasing negative pressure in the pleural cavities, is to increase the flow of blood from the veins into the heart, especially on the right side, without affecting to any appreciable degree the outflow of blood from the heart into the arteries. The inspiratory descent of the diaphragm will, moreover, tend to increase the inflow into the heart by raising the pressure in the abdomen, so that the blood is pressed out of the abdominal veins, and sucked into the heart and the thoracic veins. *During* inspiration, therefore, the inflow into and output from the right heart will be more rapid than during expiration, and this factor in itself will tend to raise the arterial blood pressure. During expiration, all these effects are reversed.

Although extreme expansion of the lungs increases the resistance to the flow through their vessels, presumably by lengthening and narrowing the latter, expansion of the lungs in normal inspiration considerably distends the vessels, and so increases the capacity and lowers the resistance of the pulmonary circuit.* The increased circulation rate produced by inspiration also distends the vascular bed. When the thorax expands, the immediate effect of the pull on the vessels that have thin walls is to fill them up with blood, and this blood being checked in its onward movement leads at first to a reduced flow into the left ventricle. Hence, at the onset of inspiration, the systemic blood pressure falls even further than during expiration. For the converse reason that the expanded vessels partly empty at expiration, and add the blood so expelled to the pulmonary current, the systemic pressure at the commencement of expiration continues still further the rise established during inspiration. Lewis * regarded the pericardial pressure, *i.e.* the direct influence of the thoracic movements on the heart, as playing a much more important part than changes in the pulmonary circulation in the production of the respiratory undulations in the blood pressure. He showed, moreover, that, in man, the effect of respiration on arterial blood pressure may vary according to the type of respiratory movement, a deep costal inspiration (not prolonged) causing a pure fall, while a deep diaphragmatic inspiration gives a pure rise of blood pressure. In expiration the reverse effects hold. He concluded that, in man, it is not possible to make any *general* statement as to the nature of the blood pressure response to a particular respiratory act.

---

* DALY. *Journ. Physiol.*, 1928, **65**, 422; HEINBECKER. *Am. J. Physiol.*, 1927, **31**, 170.

For Reference

DALY (1933). Reactions of the Pulmonary and Bronchial Blood Vessels. *Physiol. Rev.*, **13,** 149.

WIGGERS (1921). The Regulation of the Pulmonary Circulation. *Physiol. Rev.*, I. 239.

THE HEPATIC CIRCULATION

The blood flow through the liver is very large, about 1.5/min. in an adult man. The liver is supplied by two main vessels, the hepatic artery and the portal vein, and the former probably contributes at least one-quarter of the blood flow and about 40 per cent. of the oxygen to the liver.†

The liver, on section, is seen to be divided off into lobules composed of columns of polygonal cells, radiating from the centre. The portal vein, which drains the blood from the spleen and alimentary canal, breaks up into branches which lie at the periphery of the lobules, forming the interlobular veins, and send off numerous sinusoids which pass inwards between the columns of cells to join the intralobular vein lying at the centre of the lobule. From the intralobular vein the blood passes by the large sublobular veins into the hepatic veins and inferior vena cava. In an injected specimen it is easy to see that every liver cell is adjacent to at least one blood capillary. The portal vein conveys only venous blood to the liver. The lymphatics run in Glisson's capsule.‡

The hepatic artery principally supplies the connective tissue of Glisson's capsule, in which it breaks up into capillaries, and these are drained partly into venules which join the portal vein branches and partly into the sinusoids near to the margins of the liver lobules,§ so that the whole of the venous blood derived from the hepatic artery supply mingles with the general portal blood stream.

Complete loss of its arterial supply leads to necrosis of the liver,‖ though owing to the existence of anastomotic connections between the hepatic and diaphragmatic arteries, occlusion of the main hepatic artery may not be incompatible with life, even in man.

Complete and sudden obstruction of the portal vein leads to speedy death, with engorgement of the abdominal viscera which drain into the portal vein, while the occlusion of one of its branches leads to atrophy of the lobes of liver supplied by it.

The resistance to the flow of blood through the liver is normally very slight, as is shown by the fact that the pressure in the portal vein, which drives so large a blood flow through the liver, is only about 8 mm. Hg. The pressure shows respiratory fluctuations, rising at inspiration, owing to a massaging action of the rhythmically descending diaphragm.

---

[*] LEWIS. *J. Physiol.*, 1908, **37,** 213, 233.

[†] BURTON-OPITZ. *Q. J. Exp. Phys.*, 1911, **4,** 113; GRAB, JANSSEN and REIN. *Z. f. Biol.*, 1929, **89,** 324; SCHWIEGK. *Arch. f. exp. Path. Pharm.*, 1932, **168,** 694.

[‡] LEE. *Contrib. to Embryol.*, 1923, **15,** 63.

[§] OLDS and STAFFORD. *Bull. Johns Hopk. Hosp.*, 1930, **47,** 176.

[‖] BAINBRIDGE and LEATHES. *Biochem. J.*, 1907, **2,** 25; CAMERON and MAYES. *J. Path. and Bact.*, 1930, **33,** 799.

When there is a permanent increase of the resistance, as happens in diseases attended by portal cirrhosis, the blood flow along the portal vein finds its way into the inferior vena cava partly through a short circuiting of the liver by the enlargement of anastomotic connections which normally are insignificant, *e.g.* between the gastric and the œsophageal veins, the inferior mesenteric and the hæmorrhoidal branches of the internal iliac veins or between the coronary veins of the stomach and branches of the phrenic veins: the spleen becomes enlarged and the surface of the abdomen is covered with the enlarged veins subserving the anastomosis. It is possible to create such a condition experimentally in animals by constricting the portal vein,* and use has been made of this fact as a preliminary to extirpating the livers of animals for experimental purposes.

It is impossible to remove the liver directly in mammals, since ligature of the portal vein, which would be a necessary step in the operation, causes the blood to be dammed up in the portal area, and the animal dies within a few hours, being bled to death, so to speak, into its portal vessels. A way of partially obviating this difficulty was suggested by Eck.† Before ligature of the portal vein, this vessel was joined side by side to the vena cava, so that after the ligature the portal blood could flow directly into the general circulation without passing through the liver (Eck fistula). The liver was not removed, and a path to it was still open by the hepatic artery. As liver extirpations, the experiments were inconclusive, because the liver was still present and provided with a large, though abnormal, circulation. The operation still has its uses, however, in cases where it is desired to short circuit the portal circulation.

More satisfactory extirpations were performed by Mann and Magath ‡ and their collaborators. The operation is performed in three stages. First an operation like an Eck fistula is performed, except that after connecting the portal vein laterally with the vena cava, the vena cava and not the portal vein, is tied above the anastomosis. All the blood from the vena cava below the liver is thus diverted into the portal vein. After recovery from this it is found that owing to the increased pressure in the portal vein a bye-pass circulation is soon established between the portal vein and thoracic veins, *e.g.* the internal mammary and azygos. When this has occurred, the second operation, viz. ligature of the portal vein at the portal fissure, is performed. Now all the portal blood flows by the collaterals to reach the inferior vena cava above the liver. Lastly, the hepatic artery is ligatured and the liver removed, together with the length of inferior cava into which its vessels drained.

The portal area also plays an important part in the circulation by acting as a variable reservoir for blood.§ Not only is the volume of blood which passes the liver great, but the channels through which it is conducted are wide, so that the amount of blood present in the liver at any moment is considerable. The portal vein, with its rootlets (including the spleen), represents a reservoir intercalated between two resistances, viz. the arterioles of the splanchnic area centrally and the small liver vessels peripherally. This reservoir when distended can hold about a third of the blood in the body. Filling of the portal reservoir can be effected by dilatation of the primary (splanchnic) resistance or by constriction of the secondary (hepatic) resistance, and conversely a partial emptying of the reservoir results from constriction of the primary and/or dilatation of the secondary resistance.

The regulation of the splanchnic arterioles, by which the rate of entry of blood into the portal vein is controlled, is by the usual vasoconstrictor nerves of the splanchnic area. On stimulation of the splanchnic nerves the portal vein pressure, which is usually low (about 8 to 10 mm. Hg), at first falls a little and then rises abruptly ‖ to about double the normal value. The cause of the small and transitory initial fall is probably the immediate reduction of the flow of blood into the portal vein in consequence of constriction of the splanchnic arterioles. The subsequent rise is due to several causes, and is no doubt similar to the rise produced by adrenaline.¶ Part of it is due to constriction of the smallest branches of the portal vein within the liver. There is, however, another mechanism of considerable importance, viz. the communication, in the liver, between the hepatic artery and portal vein circulations. In consequence of this, the rise of general arterial pressure, which results when the splanchnics are stimulated, or when adrenaline is given, causes more blood to pass to the liver by the hepatic artery and so to raise the pressure in, and thus act as a resistance to the flow through, the vessels supplied by the portal vein.

---

* Cl. Bernard. "Le diabète," 1877, p. 317. (Paris, Baillière.)

† Dale and Laidlaw, *J. Physiol.*, 1919, **52**, 351 (simplified method).

‡ Mann, *Amer. J. Physiol.*, 1921, **55**, 285; Mann and Magath, *Ergebn. d. Physiol.*, 1924, **23**, 212 (simplified method); Soskin, *J. Lab. Clin. Med.*, 1931, **16**, 382.

§ Krogh. *Skand. Arch. f. Physiol.*, 1912, **27**, 227.

‖ Bayliss and Starling. *J. Physiol.*, 1894, **16**, 170; 1895, **17**, 120.

¶ McMichael. *J. Physiol.*, 1932, **75**, 241.

That there is a connection between the hepatic artery and portal vein circulations is shown by the further fact that when the superior mesenteric artery or vein is tied off, the portal vein pressure, though it falls, does not fall to a low level. Clamping of the hepatic artery alone also causes a fall of portal pressure for the same reason. A further factor which causes an increase of portal pressure after the administration of adrenaline is that the moderate constriction of the mesenteric arterioles is afterwards overcome by the high arterial pressure,[*] so that more instead of less blood enters the portal vein.

When there is general vasoconstriction of the splanchnic area and of the liver vessels, as occurs after administrations of vasopressin, the portal pressure falls, owing to restriction of inflow, since the arterial pressure is unable to overcome the high resistance of the constricted vessels.

In some animals, notably the dog, it is found that on perfusion of the liver outside the body, the flow soon declines unless a very high pressure is used for the perfusion. This is because there is, at the outlet of the hepatic veins, a strong layer of plain muscle which can contract and so control the outflow from the liver. It is caused to contract by histamine and relaxed by adrenaline or sympathetic stimulation,[†] and by oxytocin.[‡] Under conditions of artificial perfusion, and in certain shock-like states, these sphincters are tightly contracted, with the result that the portal pressure rises and the liver rapidly swells.

THE CEREBRAL CIRCULATION

The grey matter of the brain is very richly supplied with blood vessels from the *pia mater*. For anatomical reasons it is difficult to determine the blood flow in the brain, but it has been accomplished, both directly [§] and indirectly [||] in animals, and indirectly in man.[¶] The indirect method is by an application of the Fick principle. A mixture containing 15 per cent. nitrous oxide, and 21 per cent. oxygen is given to breathe for about 10 minutes ; simultaneous samples of arterial blood and of blood from the right internal jugular vein are drawn at intervals during 10 minutes, and analysed for N_2O content. The arterial content rises more rapidly than the venous, and by suitable calculations the amount of N_2O lost per 100 c.c. of blood passing through, and the total N_2O absorbed up to a given time, can be calculated, and hence the blood flow. Such experiments show that the blood flow is normally about 54 ± 12 c.c./100 g./min., and the oxygen consumption 3·3 ± 0·4 c.c./100 g./min. The cerebral vascular resistance, mm. Hg/c.c. blood/100 g. brain/min. can also be measured, and is 1·6 ± 0·4. In man, the A—V glucose and oxygen differences between arterial and internal jugular blood are respectively about 9·5 mg. and 6·3 vols. per cent. The above data would give therefore, a blood flow of about 740 c.c./min., a glucose usage of 70 mg./min., and an oxygen usage of 45 c.c./min., for a brain of 1,360 g. weight.

The smaller arteries derived from the pial vessels penetrate into the substance of the brain and anastomose freely with one another. Capillaries are abundant, especially in the grey matter of the cerebral cortex, where they aggregate about 1 m. of capillary vessel per c.mm. of nerve tissue, a figure which, however, is low compared with, say, cardiac muscle. The white matter is less vascular and uses less oxygen than the grey.

---

[*] Clark. *J. Physiol.*, 1928, **66,** 274 ; 1930, **69,** 171.

[†] Bauer, Dale, Poulsson and Richards. *J. Physiol.*, 1932, **74,** 343.

[‡] Holtz. *J. Physiol.*, 1932, **76,** 149.

[§] Chute and Smyth. *Q. J. Exp. Physiol.*, 1939, **29,** 379.

[||] Schmidt *et al.* *Amer. J. Physiol.*, 1945, **143,** 33.

[¶] Kety and Schmidt. *Amer. J. Physiol.*, 1945, **143,** 53 ; *J. Clin. Inv.*, 1948, **27,** 376.

In all higher animals, the brain is enclosed in a rigid bony cranium. In the child, before the cranial vault is fully ossified, part of this vault consists of membrane, the anterior fontanelle. The fontanelle bulges with each heart beat, as well as with rise of venous pressure, such as that produced during strong expiratory efforts. When ossification is complete, such alterations in the volume of the cranial contents are impossible. When viewed through a skull window, little movement of the pial arteries is visible, apart from a straightening out at points where they change direction, at each heart beat. If any expansion of the arteries occurs, room for this expansion has to be found by displacement of some other part of the cranial contents. The rapidly variable contents of the cranium are the blood and the cerebrospinal fluid. Hence, if a lumbar puncture needle be inserted into the thecal space, it is seen that on occlusion of the venous return from the cranium by an inflated cuff, cerebrospinal fluid gushes out by the needle at 250–400 c.c./min., the cerebral vessels being simultaneously distended with blood at the same rate.*

Each arterial beat is associated with a corresponding compression of the veins, and if a cannula be introduced through the occipital bone into the torcular Herophili, the venous blood is seen to pulsate out with each arterial pulse. If there is a rise of arterial pressure, although the arteries may expand somewhat at the expense of the veins, there can be little dilatation of the whole organ. The first, and chief, effect of the rise of pressure will be to cause an increased pressure slope in the cranial vascular system, and, therefore, augmented velocity of flow through the system. As a result of a more prolonged rise of pressure a certain amount of dilatation of the vessels may ensue, but only at the expense of the cerebrospinal fluid. Since this is small in amount, any expansion of the brain due to vascular causes must be limited. The converse applies for shrinkage when the general arterial pressure is lowered.

BRAIN PRESSURE. If a trephine opening be made into the cranial vault, the brain bulges into the opening. By screwing a window into the trephine opening, we can find the pressure necessary to force the brain back to its previous position. This is known as the intracranial pressure ; it is, as might be expected, approximately equal to the cerebrospinal pressure and to the pressure in the venous sinuses, and is closely dependent on the latter. Forced expiratory efforts may raise the pressure from 30 to 50 mm. Hg. In the vertical position, in man, the pressure may be slightly negative, in consequence of the tendency of the venous blood to run downwards towards the heart.

Regulation of the Blood Supply to the Brain. The free anastomoses in the circle of Willis ensures an adequate distribution of blood to all parts of the brain, even should one or more of the contributory arteries get accidentally occluded. The blood supply to the brain is primarily dependent on the arterial pressure, and this, as we have seen, is specially controlled by the carotid sinuses. Hence, any tendency for the carotid arterial pressure to fall is at once counteracted by systemic circulatory changes of such a nature as to tend to restore it to normal again. If this pressure fails, the oxygen supply of the brain is affected and loss of consciousness rapidly ensues. This is what occurs when a person, who is weak from long illness, faints on suddenly getting up from bed. In the normal individual, the change in the circulation with alteration of bodily position, which would be produced by the action of gravity, is at once counteracted through the vasomotor system.

As a subsidiary means of regulation, the vessels of the *pia mater* are themselves supplied with vasomotor nerves. In order to demonstrate these it is necessary to observe the pial vessels through a window let into the skull, so as to ensure the normal rigid enclosure of the brain.† It can then be shown that stimulation of the central end of the cervical sympathetic nerve causes constriction of the pial vessels ‡ (Fig. 461). Stimulation of the vagus depressor fibres causes reflex vasodilatation, the efferent vasodilator fibres apparently reaching the cerebral vessels from the VIIth nerve.

---

* FERRIS. *Arch. Neurol. and Psych.*, 1941, **46**, 377.

† FORBES. *Arch. Neurol. and Psychiat.*, 1928, **19**, 751.

‡ FORBES and WOLF. *Ibid.*, 1928, **19**, 1057.

The action of drugs is on the whole similar to that on other vessels. Intravenous injection of small amounts of adrenaline may cause constriction, but often this is so weak that it easily gives way to passive expansion of the vessels owing to the rise of general blood pressure * (Fig. 461). Acetyl choline causes dilatation of the vessels, though if the arterial pressure falls, there may be no increase of cerebral blood flow. Histamine gives similar results; most anæsthetics also cause vasodilatation. The action of pitressin is doubtful.

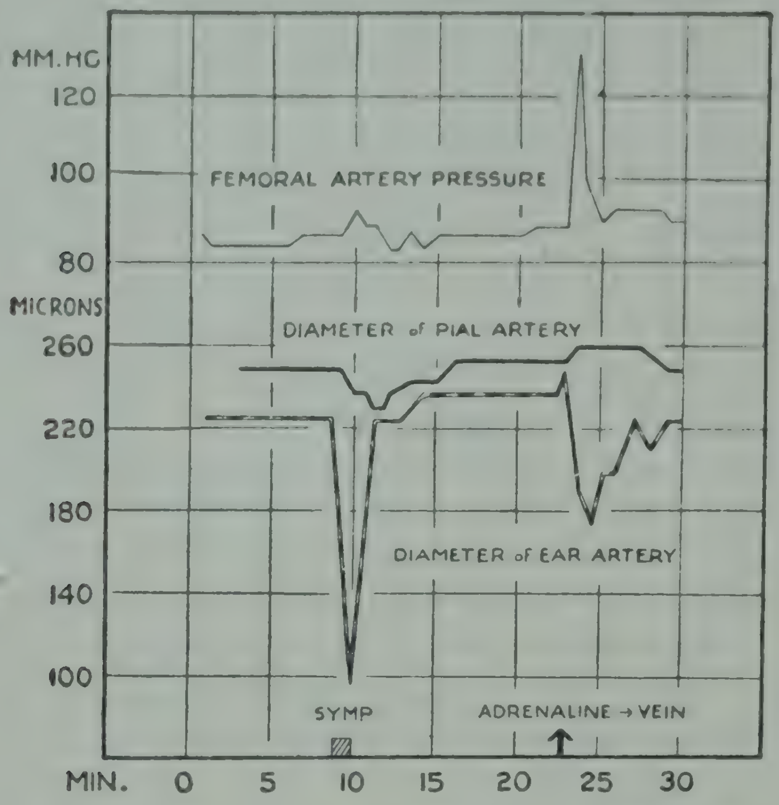


Fig. 461. Showing diameters of pial artery and of a small ear artery in the cat, together with femoral artery pressures. Observe that stimulation of the cervical sympathetic at "symp" causes much greater constriction of ear artery than of pial artery. Injection of 0·01 mg. adrenaline intravenously causes constriction of ear artery, and a large rise of arterial pressure, on account of which the pial artery passively dilates. (Forbes, Firnley and Mason.)

Under asphyxial conditions, local or general, there is dilatation of the cerebral vessels. Both oxygen lack and carbon dioxide excess act in this manner. Thus, inhaling 5–7 per cent. CO_2 may increase the cerebral blood flow 75 per cent., while 10 per cent. O_2 will increase it by 35 per cent.: increase of inhaled O_2 to 95 per cent. will reduce the flow by 13 per cent.† All these changes are due to alterations of vascular resistance.

There is some evidence that the vessels of the brain are dilated during states of increased activity.

When the intracranial pressure is increased, the arterial pressure rises, and if it exceeds the intracranial pressure, the circulation through the brain continues (or re-commences): but the resistance to the flow is raised, and so the blood flow may be reduced. When this falls below a certain level, or if the arterial pressure does not rise high enough to overcome the intracranial pressure, death results owing to asphyxia of the medullary centres.

For Reference

Symposium on Cerebral Circulation (1944). *Fed. Proc.*, **3**, 131.
Wolff (1936). The Cerebral Circulation. *Physiol. Rev.*, **16**, 545.

---

* Forbes, Firnley and Mason. *Ibid.*, 1933, **30**, 957.

† Kety and Schmidt. *J. Clin. Inv.*, 1948, **27**, 484; Courtice. *J. Physiol.*, 1941, **100**, 198.

CHAPTER XXXVI

EXAMPLES OF CIRCULATORY ADAPTATION

1. THE CIRCULATORY CHANGES DURING MUSCULAR EXERCISE

In the preceding chapters we have analysed somewhat artificially the factors which are normally involved simultaneously in the adaptation of the circulation to the necessities of the body. In order to effect a synthesis of these isolated mechanisms we may take as an example the adaptations which are involved in muscular exercise. To a large extent we can investigate the subject on man.

On initiating such experiments we meet at once with a new fact—viz. that the reflex adaptations are forestalled and reinforced by the intervention of impulses proceeding from the brain. Thus, the willed effort, or the emotion of fear or anger gives rise at the same time to impulses, starting in the brain, which excite changes in the circulatory and respiratory systems such as anticipate those which will be later excited reflexly as a result of the exercise. With a man seated on a stationary bicycle the mere question " Are you ready ? " evokes increase of muscular tone in the act of attention, increased pulmonary ventilation, and a rise of pulse rate and of blood pressure. And these changes are increased as soon as the word " go " is given. The reinforcing impulses from the cortex, which put these various mechanisms into action, are effective especially at the beginning of muscular work. The subsequent adaptations of the different organs of the body are probably chiefly automatic.

We have seen that the oxygen intake, and the CO_2 output, may undergo at least a twelve-fold augmentation during violent muscular effort carried out for a short time ; as much as a twenty-fold increase has been observed, and a five-fold increase is common and may last for many hours. A necessary condition of all muscular exercise is that the muscles shall be supplied with oxygen in proportion to their requirements. Since the arterial blood is, under normal conditions at rest, about 95 per cent. saturated with oxygen, such increases in the oxygen usage by the muscles must imply a great increase in the blood supplied to them, though a certain extra amount of oxygen may be gained by a more complete utilisation of the oxygen of the blood as it flows through the working tissues. This increased oxygen traffic involves, in its turn, an increase in the blood flow through the lungs and in the ventilation of the lungs. The circulation through the lungs has been investigated in the healthy man during rest and during exercise.*†‡ In Fig. 462 are shown results obtained by Means and Newburgh. It will be seen that the blood flow through the lungs, *i.e.* the output of the right ventricle per minute, and also the lung ventilation, increase in a manner almost absolutely proportional to the consumption of oxygen, and that all three increase *pari passu* with the work done per minute. How is this admirable adjustment of the activity of the heart and circulation to the oxygen needs of the muscles effected ? The output of the heart depends on the inflow into it, so that our problem is to determine the factors which increase this. At or before the onset of muscular exercise, unless this is quite moderate, there is relaxation of the muscle vessels and contraction of the splanchnic vessels, so that the increased output of blood is passed to the muscles, and later on also to the skin, rather than to the viscera ; the latter, owing to the raised arterial pressure, may receive about the same amount of blood as in rest. Other things being equal, it is evident that, since the blood moves in a circle, any circumstance which favours the passage of blood through the circuit as a whole will lead to an acceleration of circulation rate, and this will take the form of a more

---

* Krogh and Lindhard. *J. Physiol.*, 1913, **47**, 112.

† Means and Newburgh. *J. Pharmacol.*, 1915, **7**, 449.

‡ Bock *et al.* *J. Physiol.*, 1928, **66**, 136.

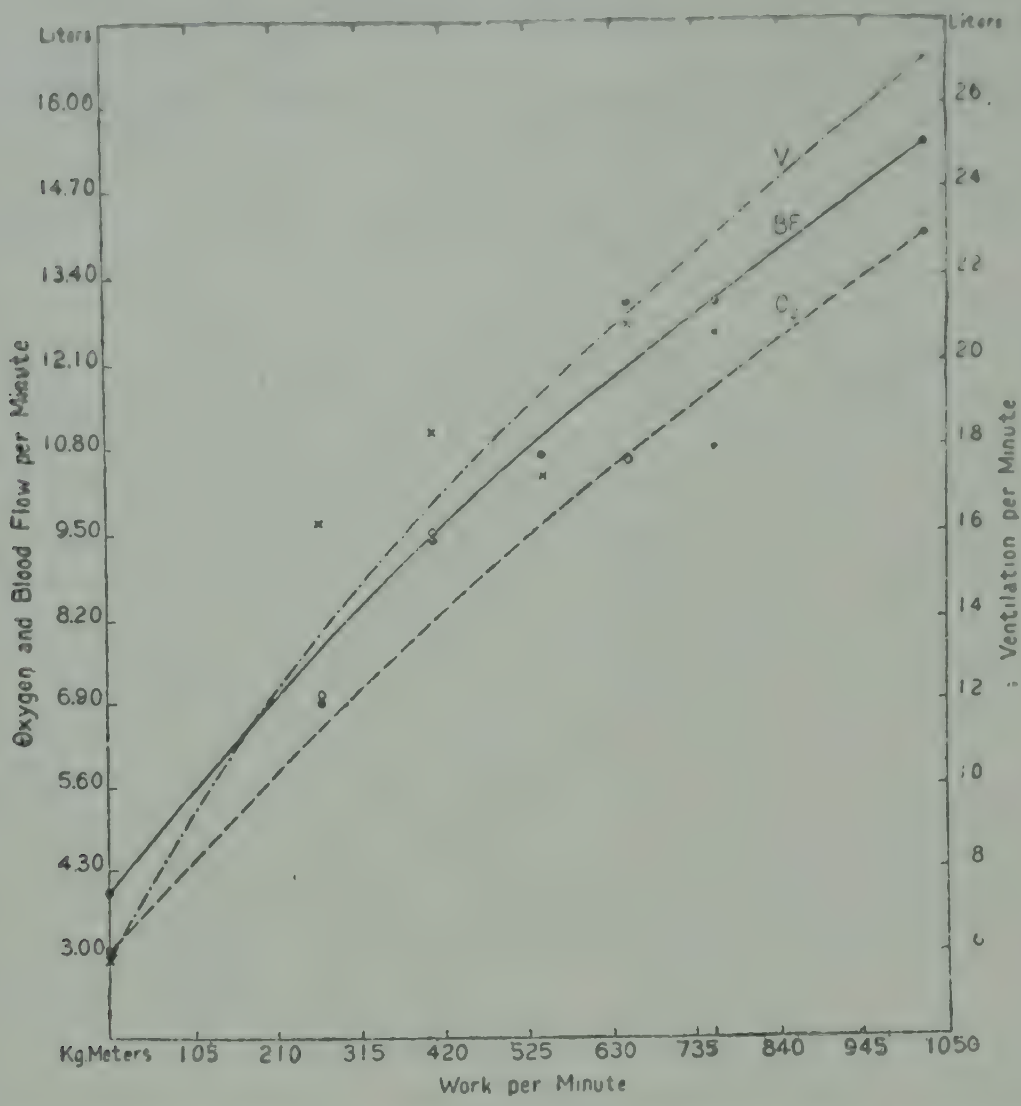
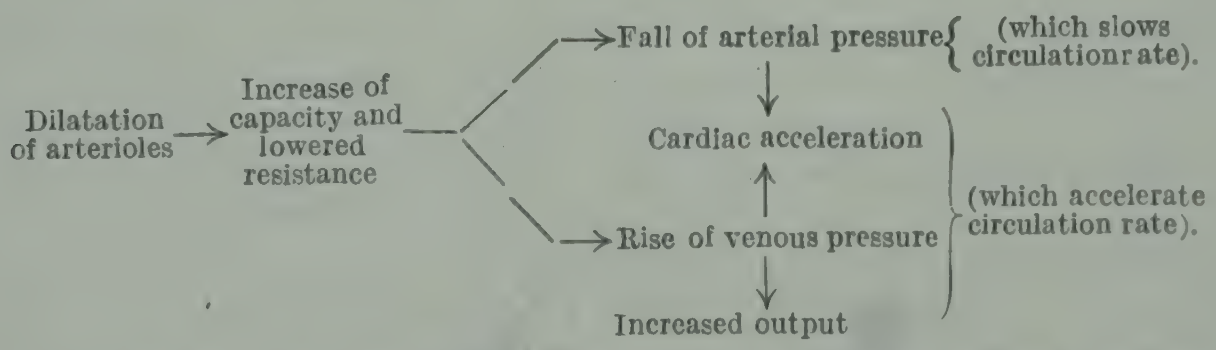


FIG. 462. Chart showing the Effect of Increasing Amounts of Muscular Work on the total Ventilation of the Lungs, V, on the Blood Flow, BF, and on the Oxygen Absorption, O_2. (From MEANS and NEWBURGH.)

rapid return of blood to the heart. But there will also be accompanying changes operating in the opposite direction, by which the result is to some extent complicated. Thus, if we imagine a simple circulation, it is clear that extensive dilatation of the arterioles will give rise to the following chain of events :


Dilatation of arterioles
Increase of capacity and lowered resistance
Fall of arterial pressure
(which slows circulation rate).
Cardiac acceleration
Rise of venous pressure
(which accelerate circulation rate).
Increased output

In consequence of such changes, a tendency to a fall of arterial pressure is more than compensated for by the accompanying alterations in the cardiac output. This is further augmented in exercise in a variety of ways. Probably the prime mover in accelerating the venous return, and so the circulation rate, is the effect of the contractions of the muscles themselves upon the rate of blood flow through their vessels. During actual contraction the flow through the muscle is slowed or even stopped.* But during and after relaxation it is agreed that the flow is increased, sometimes greatly. In the larger veins, by reason of their valves, the muscular contractions tend always to drive the blood onwards to the heart, even against a pressure of 90 mm. Hg.† The circulation through the big veins of the abdomen and chest is aided by the respiratory movements which are also augmented in proportion to muscular activity, each inspiration driving the blood out of the big veins in the abdomen and aspirating it into the thorax. The blood flow into the heart is thus increased in proportion to the activity of the muscles. The first effect of muscular exercise will be to increase the heart output at each beat. The increasing tension on the venous side of the heart evokes reflexly a quickening of the heart rhythm, chiefly by inhibition of the vagus tone, possibly also by reflex stimulation of the sympathetic accelerator nerves. Further increase in the inflow into the heart is met by corresponding quickening of the heart rhythm and the output of each ventricle per minute is increased seven to ten times. The part played by increase of output per beat, and by increase of pulse rate respectively, in augmenting the total output of the heart is shown in Fig. 463. In this figure, the first rise in pulse rate from sixty-eight to ninety-eight, and the corresponding increase in output per beat, can be regarded as associated with the initial changes originated by the act of attention and volition. It will be seen that between 270 and 600 kilogrammetres of work per minute the pulse rate remains practically unchanged, while the output per beat increases steadily with the work. After this point, there is very little further increase in the output per beat, which towards the end begins to diminish, while there is a steady increase in the pulse rate.

---

* ANREP *et al.* *Proc. R. S.*, 1934, **114B**, 223, 245 ; *J. Physiol.*, 1935, **85**, 375.

† H. BARCROFT. *J. Physiol.*, Proc. 1948, **108**, December 17th.

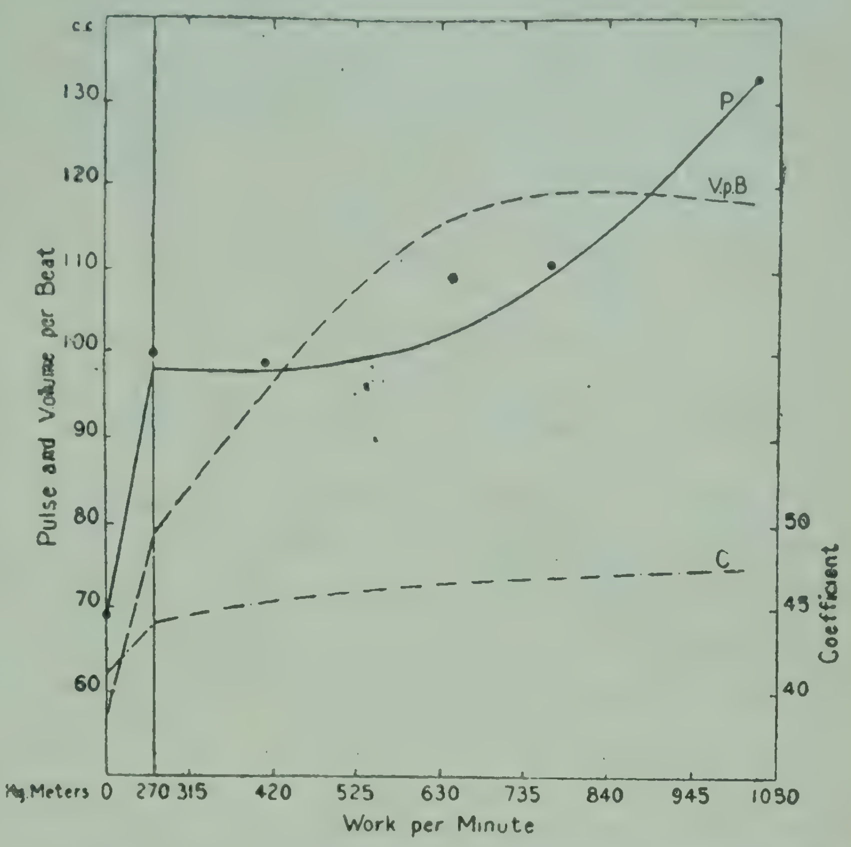


FIG. 463. Chart showing the Effect of Increasing Amounts of Muscular Work on the Pulse Rate, P, shown by dots; on the Heart Output per Beat, VpB, and on the Co-efficient of Oxygen Utilisation in the Blood, C. (From MEANS and NEWBURGH.)

By this means, the blood is driven through the lungs at a rate corresponding to the increased needs of the muscles for oxygen. The passage of this blood through the muscles is provided for by two mechanisms. In the first place, owing to the increased cardiac output, the arterial pressure is raised,* and since the muscular vessels are relaxed and the splanchnic ones constricted, a large fraction of the circulating blood can be driven through the working tissues. The total available blood is also increased by partial emptying of the portal system. In the second place, the muscles in their activity produce various metabolites, which cause further dilatation of the arterioles and capillaries in the muscles themselves. A third way in which muscular circulation is increased may be by reflexes initiated in the tissue by the presence of metabolites formed by contraction. During rest, it is probable that the majority of the capillaries are closed; during activity these dilate and are filled with blood, so that the capillary bed in the muscles may be increased many times in area, and each element of the muscle is brought into close relation with a dilated capillary, through which is flowing a rapid stream of oxygenated blood. Krogh † has shown that the number of blood-containing capillaries in each square millimetre cross-section of the muscle may be increased 40 to 100 times during maximal activity of the muscle. As a result, the oxygen tension in the muscle fibres becomes almost equal to that in the capillaries themselves.

---

* L. HILL. *J. Physiol.*, 1898, **22**, *Proc.* xxvi.; BOWEN. *Amer. J. Physiol.*, 1904, **11**, 59.
† KROGH. *J. Physiol.*, 1919, **52**, 409, 457.

Complex physico-chemical changes taking place in the blood * aid the dissociation of the oxyhæmoglobin passing through the capillaries, and therefore set free oxygen for the use of the muscles. On this account, we find almost invariably that the utilisation of the oxygen of the blood is more complete during exercise. Except during the most exhausting exercise the arterial blood is as fully saturated with oxygen as when at rest, † but in working to exhaustion it may fall to 85 per cent. saturation.‡ The fraction of the oxygen content of the arterial blood which is abstracted as it flows round the circulation is known as the "co-efficient of utilisation." Thus if 328 c.c. of oxygen were used per minute and the blood flow were 4·5 litres per minute, $\frac{328}{4\cdot5} = 73$ c.c. oxygen would be utilised per litre of blood. If the oxygen capacity of the blood were 193 c.c. per litre, the co-efficient of utilisation would be $\frac{73}{193} = 0\cdot38$. In Fig. 463 the co-efficient of the oxygen utilisation is given by the curve C. It has been found that the co-efficient of utilisation rises during severe exercise from a resting value of about 0·3—0·5 to 0·7—0·8.§

It has been shown by Cannon that every state of excitement, and especially fear and anger, is attended with increased secretion of adrenaline into the blood stream. During violent exercise associated with emotional stress, there will be an excess of adrenaline circulating in the blood, and this will reinforce the activity of the circulation. Thus, it will increase the constriction of the splanchnic area, already excited reflexly and by the central effects of the increased CO_2 or lactic acid in the blood. In the heart, the adrenaline will increase the contractile power and also the rate of beat, while, by its dilator action on the coronary vessels, it will aid the supply of oxygen to the heart

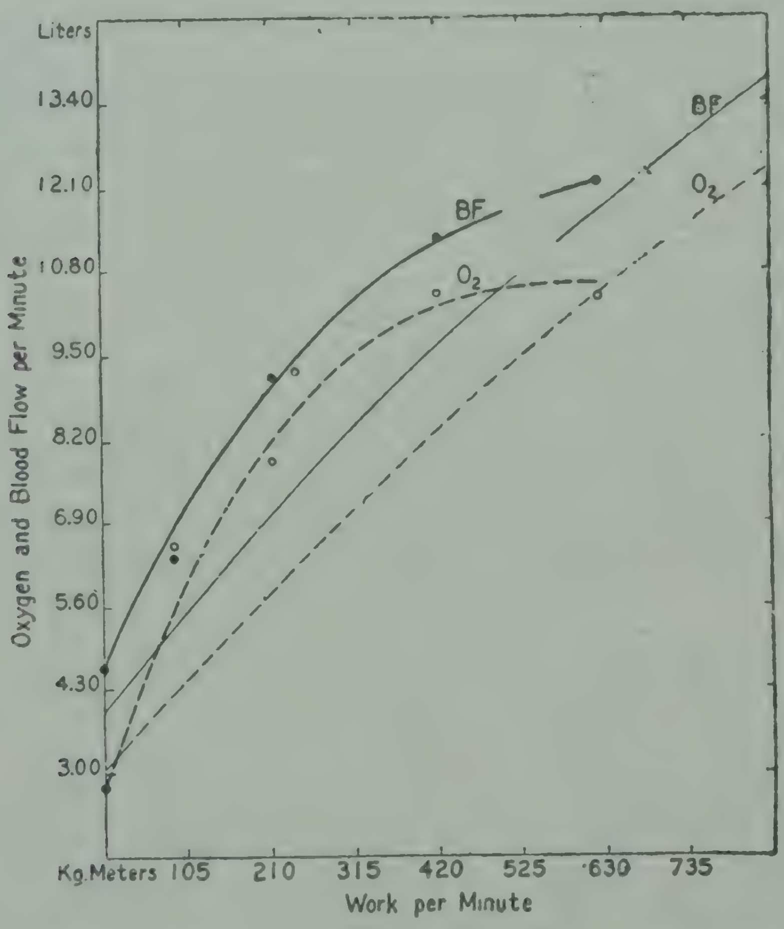


FIG. 464. Chart showing the Effects of Muscular Work on the Blood Flow and Oxygen Consumption in a subject with Aortic Disease, as compared with a normal individual (shown in lighter lines). (From MEANS and NEWBURGH.||)

(For oxygen consumption omit the decimal point and read in c.c.s.)

---

* HOCHREIN, DILL and L. J. HENDERSON. *Arch. f. exp. Path. u. Pharm.*, 1929, **143**, 129.

† HIMWICH and BARR. *J. Biol. Chem.*, 1923, **57**, 363.

‡ HARROP. *J. Exp. Med.*, 1919, **30**, 241.

§ DOUGLAS and HALDANE. *J. Physiol.*, 1922, **56**, 69.

|| MEANS and NEWBURGH. *J. Pharmacol.*, 1915, **6**, 441.

muscle. At the same time, it will cause a rapid conversion of the glycogen of the liver into sugar, so that the contracting muscles will be adequately supplied with a food which they can utilise. It is doubtful whether these effects of adrenaline are to be reckoned with, except in cases of severe emotional stress.

After the conclusion of exercise, the changes which accompanied it disappear. Arterial pressure falls quickly, and usually drops to below the normal, slowly to return. Pulse rate also drops very suddenly in those with efficient hearts, and sometimes shows a second slight rise accompanying the rapid fall of blood pressure. Oxygen usage and cardiac output drop simultaneously, and sooner than pulse rate, while pulmonary ventilation falls more slowly, and blood lactate falls most slowly.*

Training. The muscular efficiency of a man is measured by the extent to which he can call upon his body for increased efforts, *i.e.* by his margin of response. Over a moderate period of time, the individual may increase his muscular work, his respiratory exchange, and the rate of his circulation six times above that which obtains during rest. In a sudden or short-lasting effort, a much greater amount of energy can be put out, the muscles "going into debt" for the necessary oxygen. The performance of a normal individual can be increased by training, the essential features in which are graduated exercise and healthy diet, so that the muscle grows and becomes free from interstitial fat, while the fluid parts of the body and of the blood are diminished so that a larger amount of oxygen can be carried per unit volume of blood. The well-trained individual may have a margin of as much as 1,200 per cent. Disease is marked by a diminution of the margin. In Fig. 464 is given diagrammatically the response of the circulation and respiration of a man with heart disease affecting the aortic and mitral valves. This man had no discomfort and was able to do ordinary work without ill effects. On testing him on measured muscular tasks, it will be seen that, although at first he reacts like the normal individual, his margin is diminished, and when doing only 315 kilogrammetres of work per minute, the rise in the oxygen intake and in the heart output fails to keep pace with the increase in the work and also loses the parallelism which is so marked a feature in normal individuals. With increasing disease, the time would finally come when the margin was reduced to 100 per cent. or 50 per cent., so that even the act of changing from a recumbent to an erect position might be too much for the enfeebled adaptive mechanisms of the body, and the patient would have to keep to his bed. There is thus no definite dividing line between health and disease, the change from one to the other being but a progressive diminution of margin or extent of adaptation.

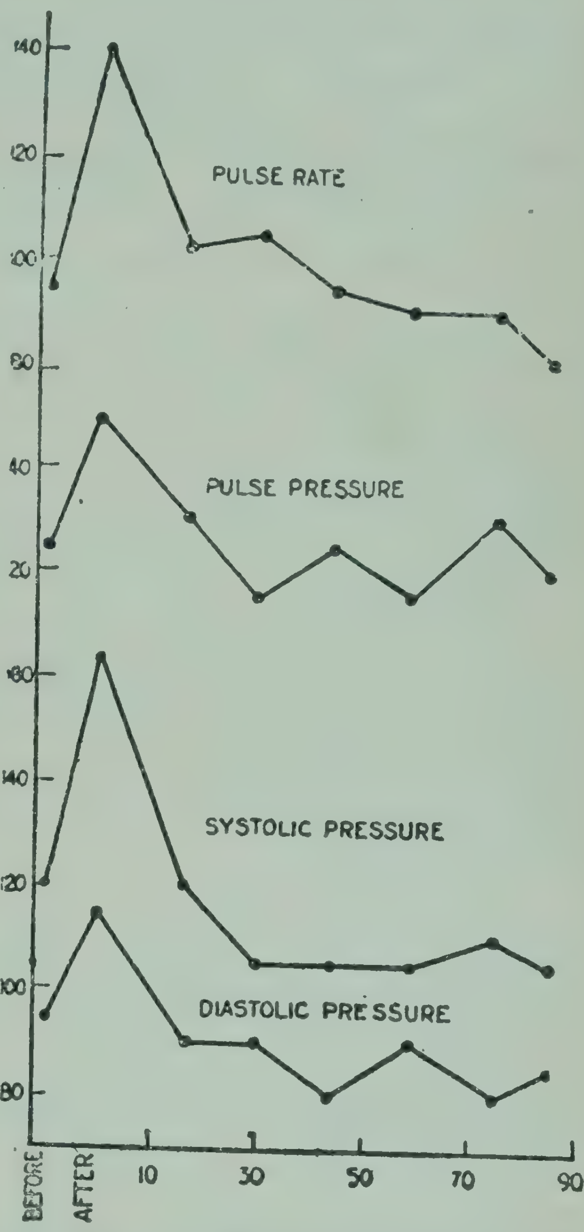


FIG. 465. Curves showing the Influence of Exercise on the Circulation. The exercise was a six-mile run. Ordinates = mm. Hg pressure and rate per minute. (O. S. LOWSLEY.†)

We have seen that the physiological condition of the heart is measured by the dilatation of its cavities, *i.e.* of the shortest length of its muscle fibres, required in order that in its beat it may set up a contractile stress adequate to expel its contents against the arterial resistance. Thus a degree of filling of the heart, which in a well-trained man may be adequate to ensure a contraction sufficient entirely to empty its cavities, in a weaker heart would be inadequate, so that blood would accumulate at each diastole until the stretching of the fibres was sufficient to ensure that the amount entering during diastole was expelled at each systole. The trained man—*i.e.* with a heart in good condition—will therefore have a considerable range over which the output per beat can be increased with increasing inflow, without alteration of rhythm. In the untrained man this margin will be smaller, so that the second mechanism of adaptation, viz. quickening of the heart beat, will be sooner brought into action to cope with the increased inflow associated with muscular exercise. Thus, one finds a considerable difference in the effect of exercise on the pulse rate in trained and untrained individuals respectively,* and this is especially shown in the promptness of recovery of the pulse when the exercise comes to an end, the effects lasting much longer in the untrained. In all cases exercise not carried to exhaustion tends to be followed by a prolonged diminution both in pulse rate and in blood pressure (*cf.* Fig. 465).

---

* BARMAN, CONSOLAZIO and MOREIRA. *Amer. J. Phys.*, 1942, **138**, 16.

† LOWSLEY. *Amer. J. Physiol.*, 1911, **27**, 446.

Ischæmic Pain. When muscles contract under such conditions that the blood supply does not increase in keeping with their increased needs, a peculiar cramp-like pain is experienced in the muscles concerned. It is met with in cases of narrowing of the arteries to a limb, and the condition is called intermittent claudication. It may be imitated by placing a sphygmomanometer armlet round the upper arm, inflating it suddenly to 250 mm. Hg, and then making gripping movements of the hand every second. The pain commences in about 35 seconds and becomes intolerable in 70 seconds; it persists as long as the occlusion of the circulation is kept up, but, on restoring it, is abolished in 3 seconds. If the circulation is again stopped, no pain is felt unless the muscles are once more exercised, but, if they are, it appears much sooner than in the first instance. These various acts led Lewis † to the conclusion that the pain is due to the accumulation during muscular contraction of a substance, which he called the " P-factor "; this under normal conditions is removed by oxidation, but when this is not done the substance stimulates pain fibres. The cardiac pain accompanying coronary occlusion is believed by Lewis to be explicable in a similar way.

For Reference

BAINBRIDGE (1931) (Edited by Bock and Dill). The Physiology of Muscular Exercise. (London: Longmans, Green.)
A. V. HILL (1927). Muscular Movement in Man. (New York and London: McGraw Co.)
SCHNEIDER (1933). The Physiology of Muscular Activity. (London: Saunders.)
STEINHAUS (1933). Chronic Effects of Exercise. *Physiol. Rev.*, **13**, 103.

2. THE INFLUENCE OF VARIATIONS IN THE TOTAL QUANTITY OF BLOOD

Plethora and Hydræmic Plethora. The effects of increasing the total volume of circulating fluid may be studied by injecting several hundred cubic centimetres of defibrinated blood or normal saline solution into a vein. The condition resulting has been called *plethora* or *hydræmic plethora* respectively. In the latter state the plasma is made more dilute and the corpuscular concentration and blood viscosity lowered. On the arterial pressure the result of such an injection is not very marked. There is a slight initial increase in the pressure, but the increase is by no means proportional to the amount of fluid injected, showing that the excess fluid is not to any large extent retained in the arterial system. On examining the pressure in the veins, however, we find a very great relative rise of pressure, and on opening the abdomen it is seen that all the veins are distended and that the liver is swollen. This is in part due to the effect of lowered viscosity of the blood, which now passes more readily through the finest vessels. The organism further prevents the rise on the arterial side by relaxing the whole system of arterioles, so that the distribution of pressures is altered, and the venous

---

[*] PEMBREY and TODD. *J. Physiol.*, 1908, **37**, *Pr.* lxvi.

[†] LEWIS. *Arch. Int. Med.*, 1932, **49**, 713.

approximates more closely to the arterial pressure (Fig. 466). This arterial dilatation and lessened viscosity augments the velocity of the blood : *e.g.* the velocity may be accelerated to six or eight times the normal rate by injecting an amount of salt solution equivalent to 50 per cent. of the total blood.

The high venous pressure causes increased diastolic filling of the ventricles, and therefore augments the output per beat. If the vagi are intact, the frequency is also reflexly raised, in consequence of the greater distension of the auricles. Thus, the work of the heart is increased in three ways—viz. by

(1) Rise of arterial pressure.
(2) Greater frequency of beat.
(3) Increased output at each beat.

These series of changes result in the relief of the vascular system. The rise of pressure in the aorta causes a reflex dilation of all the arterioles of the body, through the afferent fibres of the cardio-aortic and carotid sinus nerves. The heightened pressure in the abdominal veins and capillaries, together with the lowered colloid content of the blood, if saline is given, cause a great leakage of fluid, in the form of lymph, from the capillaries of the intestines and liver, while similar changes in the blood in the glomeruli of the kidney induce a copious secretion of urine, so that within a couple of hours after the injection of salt solution, the volume of the circulating fluid may have returned to normal.

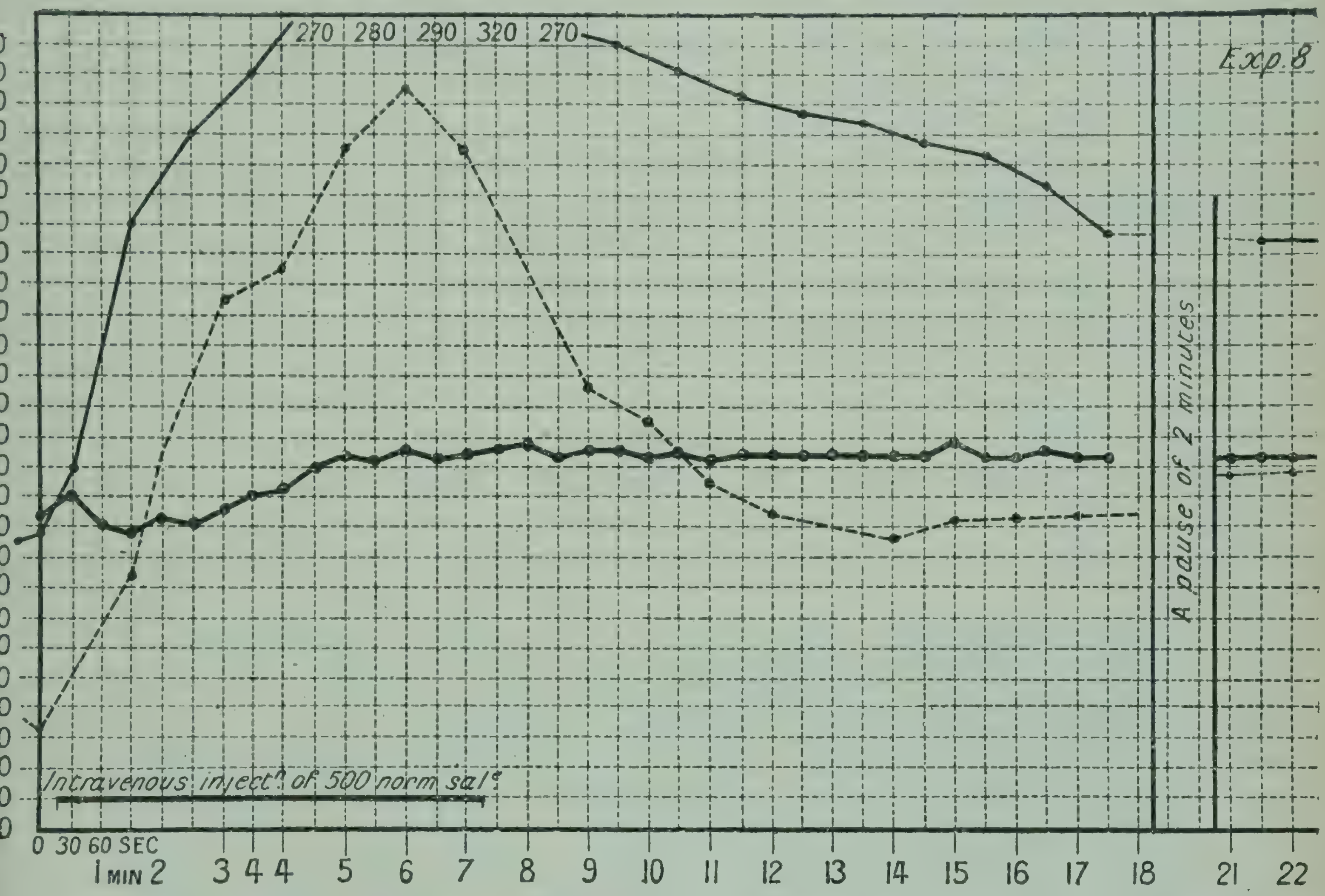


FIG. 466. Effects of Hydræmic Plethora on the Pressures in the Carotid Artery (thick line), Portal Vein (thin line), and Inferior Vena Cava (dotted line). The arterial pressure is in mm. Hg ; the venous pressures in mm. H_2O. (BAYLISS and STARLING.) *

This recovery is effected with greater difficulty if the plethora has been brought about by the injection of blood, serum, or saline containing 8 per cent. of gum arabic, since these fluids exert a colloidal osmotic pressure, and cannot escape so rapidly from the capillaries, nor be excreted unchanged by the kidneys. Hence it is easy to kill an animal by wearing out its heart, if too large quantities of blood be injected.

Hæmorrhage and Anæmia. Any diminution of the total volume of the blood, as by bleeding, would tend to lower the pressure on both sides of the system. The vasomotor centre, however, reacts to maintain the normal arterial pressure, and so the circulation through the brain, unaltered. This is attained reflexly by an excitation of the sympathetic system, which includes an outpouring of adrenaline into the blood stream. In consequence there is considerable vascular constriction, which diminishes the total capacity of the system and tends to keep the blood on the arterial side. Hence, a slight loss of blood has no influence on the arterial blood pressure, and even considerable loss may only cause a small reduction, but it causes a fall of pressure in the veins, increased respiratory movements, blanching of the skin, an increase of blood sugar, and diminished flow of urine. The heart beats more frequently, and so aids in keeping up the cardiac output.

---

* BAYLISS and STARLING. *J. Physiol.*, 1894, **16**, 159.

The compensatory vasoconstriction produced by hæmorrhage is, however, not exerted equally or indiscriminately throughout the body.* It is greatest in the skin, and next in order come the spleen, glands, and mucous membranes: the alimentary canal and organs concerned with swallowing are less affected. This is a reasonable adaptation, since if absorption were affected, as it would be if there was widespread splanchnic constriction, the lost blood could not be readily replaced.

The constriction of skin vessels after hæmorrhage persists for an hour or more after the removed blood has been replaced by transfusion. Even if more blood is restored than was lost the constriction still remains for some time, though a generalised reactive hyperæmia might have been expected.† Thus the capacity of the circulatory system is kept small and shock is warded off.

The deficiency of circulating fluid caused by bleeding is soon remedied by a transfer of fluid from the tissues to the blood. This transfer is independent of the flow of lymph from the thoracic duct into the blood, and is the direct consequence of the universal fall of capillary pressure which results from the bleeding. The abstraction of fluid from the tissues is responsible for the extreme thirst which is one result of hæmorrhage. The transfer of fluid from tissues to blood is often rapid; even during the course of a bleeding, it is found that the later samples of blood are more dilute than those obtained at the beginning. The blood also clots more rapidly. Thus the loss of fluid is replaced and further bleeding checked. After a bleeding, however, an animal has lost proteins and blood corpuscles, and these constituents of the blood are but slowly restored, the former indirectly from the food, the latter by an increased activity of the red marrow.

Physiological Alterations in the Filling of the Vascular System. The volume of blood in circulation is not quite constant, but often undergoes small and rather rapid fluctuations owing to an altered balance of loss and gain of fluid through the capillary walls. Thus, a rise of general arterial pressure, or of capillary pressure, is accompanied by concentration, and a fall by dilution, of the red corpuscles, as if to compensate. This alteration of corpuscular content is, however, sometimes in part due also to alterations in the volume of the spleen, as is explained later.

The general capacity of the vascular system is also variable. Great variations can be produced by the opening or closing of the capillary network in large masses of tissue. Among the larger vessels the most variable in capacity are the great veins; these in the human subject can hold 500 c.c. or more when distended, but usually they form baggy reservoirs of very variable content. The capacity of the venous system is very important in relation to the portal circulation, discussed in an earlier chapter. The portal vein, with its rootlets, represents a reservoir which when distended can hold about a third of the blood in the body. Partial depletion of the portal venous system is what happens, for instance, in exercise; the result is a fall of portal pressure, a rise of caval pressure, and a greatly increased venous return and cardiac output, which is due to a virtual increase in the amount of blood in the circulatory current.

---

* Rous and Gilding. *J. Exp. Med.*, 1929, **50**, 189. Gilding. *J. Exp. Med.*, 1929, **50**, 213.

† Gilding and Rous. *J. Exp. Med.*, 1929, **50**, 471.

The Spleen. The spleen may be regarded as a further special part of the portal system which, in some animals, acts especially as a reservoir for red cells. The investigations of Barcroft showed that the size of the spleen varies with physiological requirements, as such a view would anticipate. It may be extirpated without ill-effect, and occasionally is congenitally absent.

This organ is formed of a framework of connective tissue and unstriated muscular fibres, in the interstices of which is contained the *splenic pulp*. This consists of a fine fibrillar network, the meshes of which contain the cells of the splenic pulp. These are monocytes, and closely related cells giving origin to them, red cells and leucocytes. The blood from the splenic pulp is collected into large venous sinuses which run along the trabeculæ to the hilum, where they unite to form the splenic vein. These sinuses are lined by cells belonging to the reticulo-endothelial system, and called histiocytes or macrophages. They, and the monocytes of the pulp, are actively phagocytic.

The arteries to the spleen are beset in their course along the trabeculæ with small nodules of lymphoid tissue, which are known as the Malpighian follicles. These contain germinal areas in which, by division, new lymphocytes are produced, and find their way into the blood. The arteries yield a few capillaries which supply the Malpighian follicles, but for the most part open either directly into the venous sinuses, or else, by way of curious perforated sacs, into the pulp. From the pulp, the blood can pass by numerous spaces into the venous sinuses, and so into the splenic vein.

There are thus alternate paths for the circulation through the spleen; that from arteries direct to sinuses is probably free, but it is evident that the blood must meet with considerable resistance in passing through the close meshwork of the splenic pulp. This path of the circulation is, however, assisted by the occurrence of rhythmic contractions of the plain muscle forming the capsule and trabeculæ. If the spleen volume be recorded, it will be seen to be subject to a series of slow variations, each contraction and expansion lasting about a minute and recurring with great regularity (Fig. 467). In addition to these rhythmic variations in size the spleen can also alter its mean volume very greatly.

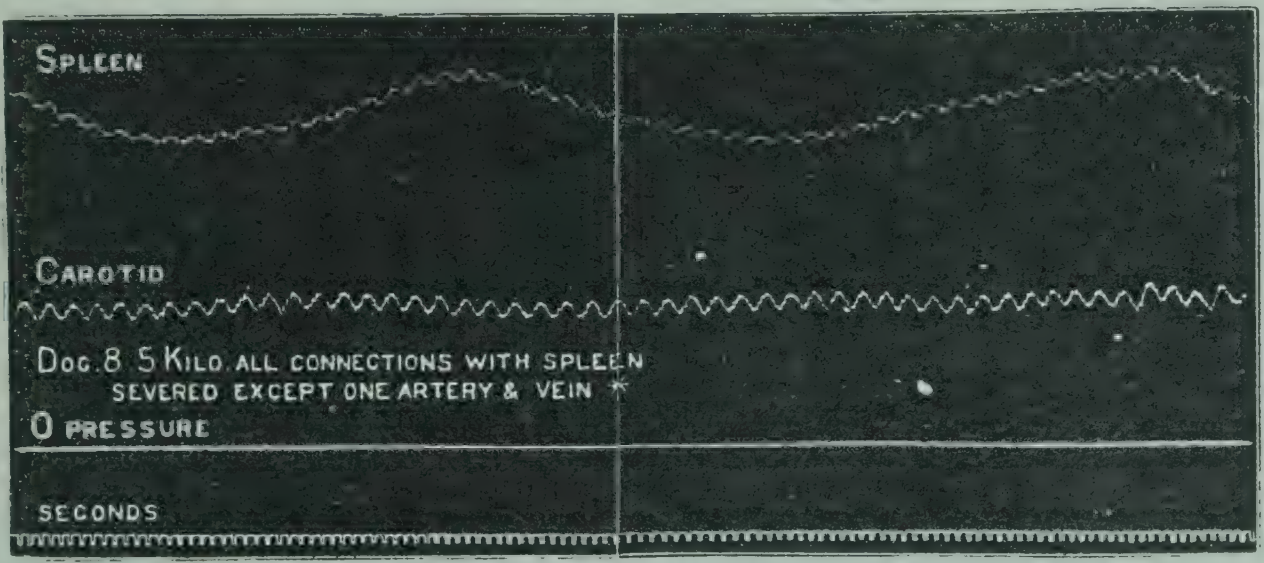


FIG. 467. Plethysmographic Tracing of Spleen (upper curve) from a Dog, showing the Spontaneous Contractions of this Organ. (Reduced from a tracing by SCHAFER.)

In spite of these rhythmic contractions, however, the normal pulp circulation is probably very sluggish. Thus when rats were slowly poisoned with carbon monoxide, it was a long time before the splenic blood, drawn by puncture, became as saturated with CO as that in the general circulation.* The alteration of the mean volume of the spleen is under the control of the nervous system, and a rapid and extensive contraction may be induced by stimulation of the splanchnic nerves,* and during asphyxia or anoxæmia.† Under these conditions, a considerable volume of blood, which may be very rich in red cells, is expressed into the general circulation.

---

* BARCROFT, J., and BARCROFT, H. *J. Physiol.*, 1923, **58**, 138.

FUNCTIONS OF THE SPLEEN. A consensus of opinion considers the spleen to have important functions in connection with the destruction of red blood cells. Erythrocytes of the splenic pulp are less resistant to hypotonic salt hæmolysis and more resistant to saponin hæmolysis than are the red cells of the general circulation.‡ After removal of the spleen the hypotonic fragility of the red cells is for a time reduced. Microscopic examination shows that the macrophages and monocytes of the spleen are full of particles of brown iron-free pigment (hæmofuscin) or fragments of red corpuscles,§ and we have already discussed the destruction of red corpuscles and the formation of bilirubin in these cells (p. 521). Part of the iron liberated from the disintegrated hæmoglobin is retained in the reticulo-endothelial cells in some form of combination with other substances (hæmosiderin). Probably most of the iron-containing substances are carried away by monocytes and detached reticulo-endothelial cells and temporarily stored in the liver. In some forms of anæmia (associated with hæmolytic jaundice) the reticulo-endothelial cells in the spleen are overabundant and hyperactive and the spleen is enlarged ; relief of the condition is effected by removal of the spleen, after which operation the number of red cells in the blood is increased and the bilirubinæmia is reduced.

In many cases of infectious disease, such as recurrent fever, the splenic phagocytic cells are observed towards the end of the attack to be full of the organism—a spirillum—which is the cause of the disease. In fact, these cells can take up foreign solid particles of all sorts held in suspension in the blood plasma. We must, indeed, look upon the spleen as the great blood filter, purifying the blood in its passage by taking up particles of foreign matter and effete red corpuscles. It also produces lymphocytes and monocytes, as we have already seen. In embryonic life, and in the adult in certain types of anæmia, it also produces leucocytes and red cells, just as the red marrow does.

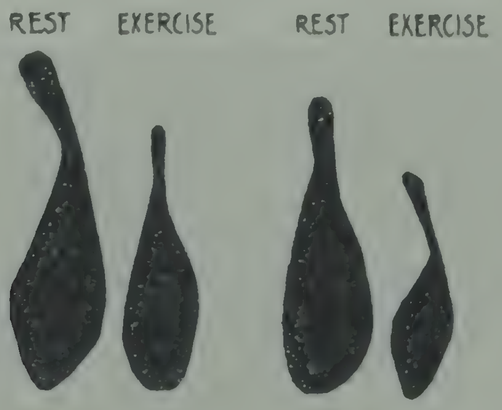


FIG. 468. Shrinkage of the Spleen of the Cat with Exercise. Reconstruction from X-ray outlines as seen from the side. (BARCROFT, *Lancet*, 1925 (I.), **208,** 319.)

Barcroft has shown that in some species the spleen may serve as a reservoir of red blood corpuscles. When there is a deficiency of hæmoglobin or of oxygen in the blood, the spleen contracts, driving some of the corpuscles it contains into the general circulation.‖ Thus it contracts in exercise, in asphyxia, after hæmorrhage, etc., to such an extent that it may be reduced to a fraction of its former size (Fig. 468). No doubt the liberation of adrenaline into the blood also greatly assists the contraction of the spleen under these conditions. After removal of the spleen, carbon monoxide is more rapidly fatal than to normal animals.* The human spleen does not store red cells in appreciable amounts.†

---

* ROY. *J. Physiol.*, 1881, **3,** 203 ; BARCROFT *et al.* *J. Physiol.*, 1932, **74,** 321.

† SCHÄFER and MOORE. *J. Physiol.*, 1896, **20,** 1.

‡ ORAHOVATS. *J. Physiol.*, 1926, **61,** 436.

§ ADDISON. *Amer. J. Anat.*, 1919, **26,** 437 ; CUNNINGHAM, SABIN and DOAN. *Proc. Soc. Exp. Biol.*, 1923, **21,** 326.

‖ BARCROFT *et al.* *J. Physiol.*, 1925, **50,** 443 ; BARCROFT and STEPHENS. *J. Physiol.*, 1927, **64,** 1.

For Reference

KRUMBHAAR (1926). Functions of the Spleen. *Physiol. Rev.*, **6**, 160.

3. THE INFLUENCE OF GRAVITY

A physical factor which may alter the pressures in the large arteries, or large veins, is that of hydrostatic pressure, due to gravity: its effects are often considerable.

As regards the arterial system, in which the pressure is normally high, the effects of gravity are relatively less than in the venous system, and, under most circumstances, the arterial pressure is able to look after itself. In the horizontal posture the mean pressures in all the larger arteries are roughly equal: in the erect posture the pressure in, say, the femoral artery is higher than that in the brachial artery by an amount equal to that of a column of blood of the height of the vertical distance between them, say about 55 cm. (= 40 mm. Hg): in the vertical head-down position there would be the same difference, but in the opposite direction. It is well known that when we rise from the recumbent to the erect posture there is a compensatory vasoconstriction and cardio-acceleration, the effect of which is to prevent the arterial pressure from suddenly falling, especially in the head vessels. We can suppose that, on standing on the head, the reverse happens. These phenomena must play a large part in regulating the arterial pressure in the cerebral arteries of a long-necked animal, such as a giraffe, in which the fluctuation in hydrostatic pressure, if not so corrected, would be of the order of at least 150 mm. Hg when the head was moved from the ground to its highest position. The arteries have thick walls, however, and except for animals like the giraffe the differences of arterial pressure between head and feet usually presents no major problem.

With the venous system it is totally different, and altogether less simple. The veins have thin walls, are often of considerable capacity, and the blood pressure within them is so low as often to be outweighed by quite moderate hydrostatic effects. If a length of a large vein, or other thin-walled elastic tube, is filled with fluid and held vertically, it tends, in consequence of hydrostatic pressure effects, to collapse above and distend below, and this tendency is inseparable from the whole of the venous system *in situ*, though it is subject to various modifications. Consequently, when we stand erect, blood tends to drop into, and accumulate (with hæmoconcentration) in, the lower parts of the body, the veins of which become distended, while those of the head and neck become empty. In the cat, and perhaps in other species too, much blood accumulates in the liver.

This tendency is offset in a variety of ways, some of them purely physical, some physiological. Among the physical factors is one referred to on p. 641, viz., that when the blood is flowing along the veins in the same direction as the pull of gravity, the collapse of the higher parts prevents the venous pressure in those parts of the vein still further up (and hence in the capillaries) from falling below a certain level. Another help is that in many parts which are normally dependent, the adventitial coat of the vein is thickened, and the walls supported by surrounding structures, so that a limit is placed to the degree of distension which may be reached. The abdominal contents, together with the tonic state of the abdominal musculature, play an important part in preventing an accumulation of blood in the capacious abdominal venous reservoirs. A third important factor is the collaboration of valves, together with muscular movement, in continually massaging blood onwards towards the heart, as already explained (p. 640).

Among the physiological factors which help to compensate the effects of gravity on the venous system, the most important is effected through the vasomotor system. As soon as blood begins to accumulate in the large veins, the venous return to the heart is reduced, and with it the cardiac output, so that the arterial pressure begins to fall, and this is aided by the direct effect of gravity on the arterial pressure. In consequence, the carotid sinus mechanism is set in operation, and there is also a reduced blood supply to the medulla. Both of these happenings result in widespread splanchnic constriction, which reduces the capacity of the circulation, and raises the cardiac output again. If this does not suffice, the respiratory centre is soon brought into a state of increased excitability, and there is increased activity of the diaphragm, the muscles of the thoracic and abdominal walls, and the levator ani, all of which by expanding the thorax and compressing the abdomen hasten the venous return.

---

* BARCROFT *et al.* *J. Physiol.*, 1925, **60**, 79.

† PARSON *et al.* *Amer. J. Physiol.*, 1948, **155**, 239.

The effect of these various factors may be shown experimentally by observing the effect of altering the position of an animal from the horizontal to the vertical head-up position. In some animals, *e.g.* snakes, or hutch-bred rabbits, compensation is so inadequate that the effect of this change is to allow the blood supply to the head to fall to so low a level that consciousness is soon lost. In most healthy mammalia, including man, compensation is usually adequate to maintain a sufficient circulation to all parts. If a normal man tries to stand perfectly still for some time, however, as with a soldier on parade, the lack of massaging action of the leg and other muscles may lead to withdrawal into the lower veins of enough blood to cause fainting : the same thing may occur when a person first stands up after having been confined to bed for some days, particularly if the abdominal muscles have become weak and flaccid.

When an animal is suddenly brought to the head-up position, the arterial blood pressure shows a sudden fall, which is normally followed by a more or less complete recovery, sometimes to above normal. In the cat, evisceration does not prevent the fall, though it reduces the subsequent compensation ; but removal of the liver does prevent the fall.* Denervation of the carotid sinuses may reduce the extent of the return of arterial pressure, while section of the lower cervical cord, or of the splanchnics, greatly reduces the power of compensation, the arterial pressure remaining at a low level as long as the vertical posture is maintained.

All the effects of gravity can be increased by the introduction of centrifugal force, and it has been shown for instance that when animals are spun in the hind-feet-outwards position in a centrifuge, they are soon killed, because the normal compensatory mechanisms are inadequate to prevent the retention of a large fraction of the blood in the legs and lower abdomen. These facts are important in high-speed aviation. Gravity normally acts in man in a head-to-feet direction, and when an aviator performs a tight inside (*i.e.* head-to-centre) turn its effect is increased by centrifugal force, in proportion to the value of $+ g$. This puts a good deal of strain on his compensatory mechanisms, and, when the force reaches a value equal to about $+ 5\ g$., the aviator in about 4 seconds first becomes temporarily blind (" black-out "), due to insufficient blood supply to the retina, and then loses consciousness.† The effect is reduced, as might be expected, if the aviator lies or crouches in a plane parallel to the direction of the turn being made, and is somewhat diminished by the wearing of a tight abdominal binder or water-filled suit.‡

When flying inverted, or performing outside (*i.e.* head-to-circumference) loops, the normal effect of gravity is reversed ($-g$.), and this, by increasing arterial pressure in the head, introduces a danger of cerebral hæmorrhage. Up to about $-5\ g$., however, there is little risk of this, because of the support provided in the rigid cranium by the cerebro-spinal fluid,§ in which pressure also rises.

* EDHOLM. *J. Physiol.*, 1942, **101**, 1.
† STEWART. *J. Physiol.*, 1945, **104**, 7.
‡ RUSHMER. *Amer. J. Physiol.*, 1947, **151**, 459.
§ RUSHMER *et al.* *Amer. J. Physiol.*, 1947, **151**, 355.

CHAPTER XXXVII

LYMPH, TISSUE FLUIDS, AND CEREBRO-SPINAL FLUID

LYMPH AND TISSUE FLUID

In all parts of the body the blood flows in capillaries with definite walls and is thus separated from the tissue-elements by these walls and by a varying thickness of tissue. In some organs, such as the liver and lung, every cell is in contact with the outer surface of some capillary; while in others, such as cartilage (which is avascular), a considerable thickness of tissue may separate any given cell from the nearest capillary. A middleman is thus needed between the blood and the tissues, and this middleman is the *tissue fluid,* which fills spaces between all the tissue elements, so that any tissue can be regarded as a sponge soaked with fluid exuded from the blood capillaries; as any substance, oxygen or foodstuff, is taken up by a tissue cell from the fluid surrounding it, this latter recoups itself at once at the expense of the blood.

Throughout these spaces we find a close network of vessels, lined and separated from the tissue spaces* by a layer of extremely thin endothelial cells, and this plexus communicates with definite channels—lymphatics, by which any excess of fluid (*lymph*) is drained off. New lymphatic vessels have been seen to grow by sprouting from previously existing lymphatics.† The lymphatics all run towards the chest, where those of the hind limbs join a large vessel (the *receptaculum chyli*), which receives the lymph from the alimentary canal, to form the *thoracic duct.* This, after receiving the *left cervical duct* from the left fore limb and the left side of the neck, opens into the venous system at the junction of the left internal jugular with the subclavian vein. A small *right lymph duct* drains the lymph from the right fore limb and right side of the chest and neck.

The lymph may be looked upon as derived from the tissue fluid which exudes from the plasma, through the capillary wall, supplies all the tissue elements, and contributes to the formation of secretions of glands; the excess of fluid (Fig. 469) passes between or through the endothelial cells into the peripheral lymphatic network, whence it is carried by lymphatic trunks into the thoracic duct, by which it is returned again to the blood.

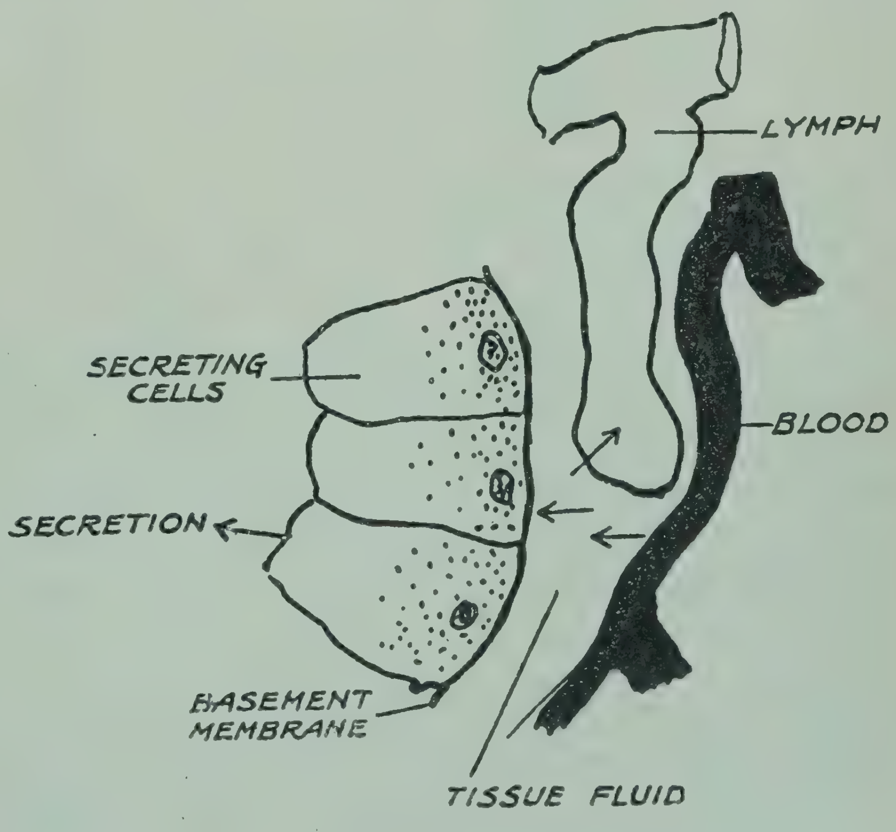


Fig. 469.

It is important, therefore, as Drinker has pointed out, always to regard the blood capillaries and the lymph system as serving together to " the same end—the maintenance of the internal environment at constant composition and volume."

---

* MacCallum. *Arch. f. Anat. u. Phys. (Anat. Abt.),* 1902, 273.
† Clark and Clark. *Amer. J. Anat.,* 1938, **62,** 59.

It is easy to obtain lymph for examination by putting a cannula into the thoracic duct, and collecting the fluid that drops from it. We may also tap in a similar way the lymphatics from the heart,* or lungs, or one of the large lymphatic trunks of the limbs †; but in the latter case we have to move the limb to induce a flow of lymph.

Properties of Lymph. Lymph obtained from the thoracic duct varies in composition and appearance. From a fasting animal the lymph is a transparent, slightly yellowish liquid. When obtained from an animal shortly after a fatty meal, it is milky from the presence of minute particles of fat absorbed from the alimentary canal, and, if the intestines be exposed, the small lymphatics are to be seen as white lines running from the intestine to the attached part of the mesentery. It is owing to this fact that these lymphatics have received the special name *lacteals*, the lymph in them, when milky, being called *chyle.*

On microscopic examination, the lymph presents lymphocytes similar to those of blood, which are, indeed, partly derived from the lymph. All the lymphatics pass at some point of their course through lymphatic glands, which add lymphocytes. Lymphocytes are also formed in all localities where we find adenoid tissue, such as the tonsils, air passages, Peyer's patches and solitary follicles, Malpighian bodies of the spleen, and thymus.

The lymph from the thoracic duct clots slowly after it has left the vessels, forming a colourless clot of fibrin. It contains about 6 per cent. of solids which are similar to those of the plasma, and are present in similar proportions except that it contains much less protein, as the following table shows :—

COMPARISON OF SERUM AND LYMPH OF DOG. (ARNOLD and MENDEL, *J. Biol. Chem.*, 1927, **72**, 189.)

	Serum g./100 c.c.	Lymph g./100 c.c.
Total solids . .	8·3	5·2
Chlorides . .	0·392	0·413
Calcium. . .	0·0104	0·0093
Phosphorus . .	0·0043	0·0036
Glucose . .	0·123	0·124
Non-protein N. .	0·0272	0·0270
Protein N. . .	0·90	0·57

Anything causing a change in composition of the blood, *e.g.* feeding with glucose, injection of insulin or parathyroid hormone, removal of the kidneys or parathyroids, etc., causes a comparable change in the composition of the lymph.‡ Liver lymph may contain more sugar than that from other parts if hepatic vein blood contains more sugar than portal vein blood.§

The Production of Lymph. The total volume of lymph in the body is no doubt variable from time to time, but is probably of about the same order as the blood volume. The volume of tissue fluid is probably similar under normal conditions, though much increased when tissues become œdematous. Investigation has shown that in the formation of tissue fluid and lymph the endothelial cells act passively, as colloidal membranes of variable permeability.

A comparison of the composition of serum, heart lymph and pericardial fluid shows that the last two resemble filtrates from the blood :—

---

* DRINKER *et al.* *Am. J. Physiol.*, 1940, **130**, 43.

† WHITE, FIELD and DRINKER. *Amer. J. Physiol.*, 1933, **103**, 34.

‡ HEIM, THOMSON and BARTTER. *Amer. J. Physiol.*, 1935, **113**, 548.

§ CLARK and WINTER. *J. Physiol.*, 1935, **83**, 49P.

SERUM, HEART LYMPH AND PERICARDIAL FLUID. (In round figures, from DRINKER *et al.*, 1940.)

Fluid	Total protein	Albumin/ globulin ratio	Chloride	Colloid osmotic pressure mm. H_2O	
			mg./100 c.c.		per g. protein
Serum . . .	5·95	1·01	416	234	38
Heart lymph . .	3·83	1·34	450	175	44
Pericardial fluid . .	1·63	1·14	434	59	43

The factors which are responsible for the transudation of tissue fluid may be divided into two classes—mechanical and chemical, the former depending largely on the pressure of the blood in, and the degree of active dilatation of, the vessels, and the latter chiefly on the metabolism of the cells outside the vessels.

The amount of tissue fluid formed in any given capillary area must be dependent on the capillary permeability, and on the difference of pressure between the blood in the vessels and the fluid in the tissue spaces. This latter pressure is normally extremely low (about 20 mm. H_2O),* so that we should expect that the exudation will rise and fall as the capillary pressure is increased or diminished. The chief factor is the permeability of the walls of the blood (and lymph) capillaries.† The capillary walls in the limbs and connective tissues present a considerable resistance to the filtration of lymph, and keep back the larger portion of the proteins of the blood plasma, but the intestinal capillaries are much more permeable, and in the liver we find the greatest permeability. The following table shows how variable the flow and protein content of the lymph is in different parts of the body :—

AVERAGE FLOW AND PROTEIN CONTENT OF LYMPH IN DOGS. (COURTICE, 1943.)

	Flow, c.c./hr.	Protein g./100 c.c.
Hind leg duct	1·5	1·5
Cervical duct	1·4	2·6
R. lymph duct	2·6	3·6
Thoracic duct	26·0	4·1
Blood plasma	—	5·4

As regards the interchanges of fluid and non-colloidal solutes between blood stream and tissue fluid, it seems clear that there is a sort of circulation of fluid constantly going on. The blood pressure in a capillary falls off sharply from its arterial to its venous end. Often it may be 30 to 40 mm. Hg. at the former and only 10 to 15 at the latter. At the higher pressure, which is above that of the colloidal osmotic pressure of the plasma, fluid will be passed out by ultrafiltration ; at the further end the pressure is below that of the colloid osmotic pressure of the plasma, and so at the venous end fluid will be reabsorbed into the blood stream.

If we regard the lymph as being completely separated from the tissue fluid by the delicate walls of the lymphatic capillaries, the lymph flow will similarly depend on the relation between pressure of tissue fluid and permeability of lymphatic capillary. It has been the usual practice to consider the permeability of blood and of lymph capillaries together. The justifiability of this may, however, be doubted, since there is reason to believe that the tissue fluid, even under normal conditions, and certainly under pathological ones, can vary considerably without there being any change in the rate of lymph formation.

The trivial lymph flow from a limb lymphatic is practically unaltered by changes in the arterial supply, although an increase may be obtained by active movement or by massage, or by ligaturing all the veins of the limb so as to cause a great rise of capillary pressure. It has been shown that when a person stands still in the erect posture, there is a loss of plasma, an increase in the red cell count, and a slow swelling of the legs. This indicates that fluid has passed out from the blood into the tissue spaces of the legs, the actual loss of blood volume being of the order of 300 or 400 c.c. in half an hour or less. On resuming the recumbent posture the process is reversed.

---

* WELLS *et al.* *J. Clin. Inv.*, 1938, **17**, 489.

† STARLING. *J. Physiol.*, 1894, **16**, 224.

The lymph flow from the intestines can be measured by collecting the lymph from the thoracic duct after tying off the lymphatics which leave the liver. It will be found that lowering of the capillary pressure in the intestines, by obstructing the thoracic aorta, stops the flow of lymph, whereas a rise of capillary pressure, such as that produced by ligature of the portal vein, causes a four or five fold increase of the lymph. The lymph from the intestines is relatively rich in proteins.

The effect of rise of capillary pressure on the lymph flow is still more striking in the case of the liver. If the inferior vena cava be obstructed just above the opening of the hepatic veins, there is a rise of pressure in the liver capillaries to three or four times the normal height, and a large increase in the lymph flow from the thoracic duct. The lymph may be increased eight to ten times in amount, and it contains more protein than before.*

It is possible to prepare artificial membranes, *e.g.* of collodion, which show similar gradations of permeability ; some so slightly permeable that no colloids can be forced through them (*cf.* ultrafiltration, p. 54), others allowing some colloids, *e.g.* serum albumin (of small size of particles), to pass, while retaining others, *e.g.* serum globulin, and some even allowing nearly all the colloids of the plasma to pass. When such a collodion membrane is used in an osmometer, with blood plasma inside and water or salt solution outside, the salts and other materials to which it is permeable will pass through, and the only osmotic pressure exerted will be that of the colloids which cannot permeate. Hence the higher the permeability of the membrane, the lower the effective osmotic pressure of the colloids, and the greater the amount of plasma constituents that will pass through. If the pressure inside the osmometer is raised above the osmotic pressure of the retained colloid, fluid and dissolved substances will pass out at a rate proportional to the raised pressure.

In the case of the limbs and connective tissues generally, the second factor, that of permeability, comes more into prominence. The osmotic interchanges between blood and cell, through the intermediation of the tissue fluid, are constantly going on in the normal life of the tissue, and are quite independent of the amount of fluid produced. Thus a gland cell may use up oxygen, calcium, or sugar, and create a deficit of these substances in the layer of tissue fluid immediately surrounding the cell. There is at once a disturbance of the equilibrium, and a flow of these substances from blood to tissue fluid is set up. Such changes can occur with great rapidity. We find, for instance, that if a very large amount (40 grammes) of glucose be injected into the circulation, osmotic equilibrium between blood and tissue fluid is established within half a minute of the termination of the injection. In this case, the rise of osmotic pressure caused by the injection of the sugar at once attracts water from the tissue fluid, and this in its turn from the tissue cells, while at the same time glucose is diffusing outwards from the blood, until finally the osmotic pressure inside and outside the vessels is the same. A process of this character would more usually work under normal circumstances in the reverse direction, and lead to a passage of fluid from blood to tissues and tissue spaces, and changes of this nature have often been believed to be associated with activity of the tissue cells. Every active contraction of a muscle, for instance, is attended by the breaking down of a few large molecules into a number of smaller ones, and this causes a rise of osmotic pressure in the muscle fibre and surrounding fluid, and therefore a passage of fluid from blood to tissue. It is certain, however, that much of the increased lymph flow from a limb during contraction of its muscles is due to the mechanical effect of the movement in accelerating lymph flow, so that the factors involved are complex.

---

* STARLING. *J. Physiol.*, 1895, **19**, 15.

The lymph flow from the heart shows a very close relationship to the amount of work the organ is doing: if that is increased the lymph flow is accelerated, and the same happens if the oxygen saturation of the circulating blood is reduced. The evidence of experiment on secreting glands is conflicting, for while several early workers claimed to have obtained an increase of lymph flow from the submaxillary gland when caused to secrete by chorda stimulation, others, more recently * have found no change in the lymph flow. This might be due to the fact that, although an increased amount of tissue fluid is exuded from the capillaries, the excess is all passed on into the secretion; against this explanation, however, is the further fact that if we paralyse the gland cells by means of atropine, the active flushing of the gland, which accompanies activity, produces little, if any, increase in the lymph flow from the gland.

When we speak of the permeability of the blood "capillaries" we should strictly include the smallest arterioles and venules with them. In muscle, where observations are easily made, the walls of the arterioles have been shown † to be slightly permeable, and those of the venules easily permeable, to molecules of dyestuffs which are probably larger than glucose molecules. In the capillaries proper, there is evidence that the permeability increases, as the capillary pressure declines, in passing from arteriole to venule. When we remember that the muscle capillary is a long one, the advantage of this graded permeability is evident, for if the vessel were equally permeable throughout its length, the tissue near the venous end of it would have less opportunity to receive nutrient material or get rid of its waste products, because all the advantage of first choice would lie with the tissue at the arterial end.

It is probable that the permeability of the endothelium is also increased when there is active dilatation of the capillaries, as happens when an organ becomes active. Further, the relative insusceptibility of the limb capillaries to pressure holds only for the absolutely normal capillary. Lowering of the oxygen pressure of the blood increases the lymph flow from almost any organ, *e.g.* heart, lung or limb. Any factor which leads to impaired nutrition of the vascular wall, such as the presence of poisons in the blood or in the surrounding tissues, scalding or freezing, increases its permeability. Under such conditions, the limb capillary reacts to changes of pressure like a liver capillary, the slightest increase of pressure causing an appreciable increase in the production of tissue fluid. As soon as the permeability of the capillary wall is increased to such an extent that the plasma proteins pass through, then the osmotic pressure of the plasma proteins ceases to be an effective force working against exudation, which then continues even at the lowest capillary pressures. This increased production may be too great to be carried off by the lymphatic channels, so that the exuded fluid stays in the tissue spaces, distending them, and causing the condition known as *œdema* or dropsy. In the lungs, *e.g.* after capillary damage by inhaled phosgene, the capillaries become freely permeable and œdema fluid of high protein content fills up the alveoli and causes death.

Lymphagogues.‡ Among the substances which have a direct action on the vessel wall are those (*e.g.* extract of mussels, commercial peptone, etc.) called by Heidenhain lymphagogues of the first class. These, on injection into the blood stream, cause an increased flow of lymph from the thoracic duct. This is probably due simply to their deleterious action on the capillary wall, chiefly of the liver capillaries. In fact, the injection or ingestion of these bodies often gives rise to a copious eruption of nettlerash due to an increased exudation of tissue fluid into the meshes of the cutis. Histamine, which is present in most tissue extracts, has a similar effect.

---

* Leigh. *Amer. J. Physiol.*, 1935, **112**, 657.

† Rous, Gilding and Smith. *J. Exp. Med.*, 1930, **51**, 807.

‡ Starling. *J. Physiol.*, 1894, **17**, 30.

Lymphagogue action of the second class may be produced by the injection of large amounts of crystalloids, such as sodium chloride, into the circulation. In this case the lymph becomes much more dilute. The explanation of the action of these bodies is that after the injection a condition of hydræmic plethora is brought about. This increase in the total volume of the blood has two consequences; first, the colloidal osmotic pressure of the plasma proteins is lowered owing to dilution, and second, there is an increased capillary pressure (Fig. 470, A). That the action of these bodies is purely mechanical is shown by the fact that, if the rise of capillary pressure be prevented by bleeding the animal immediately before the injection, the increase in the lymph flow is also prevented (Fig. 470, B).

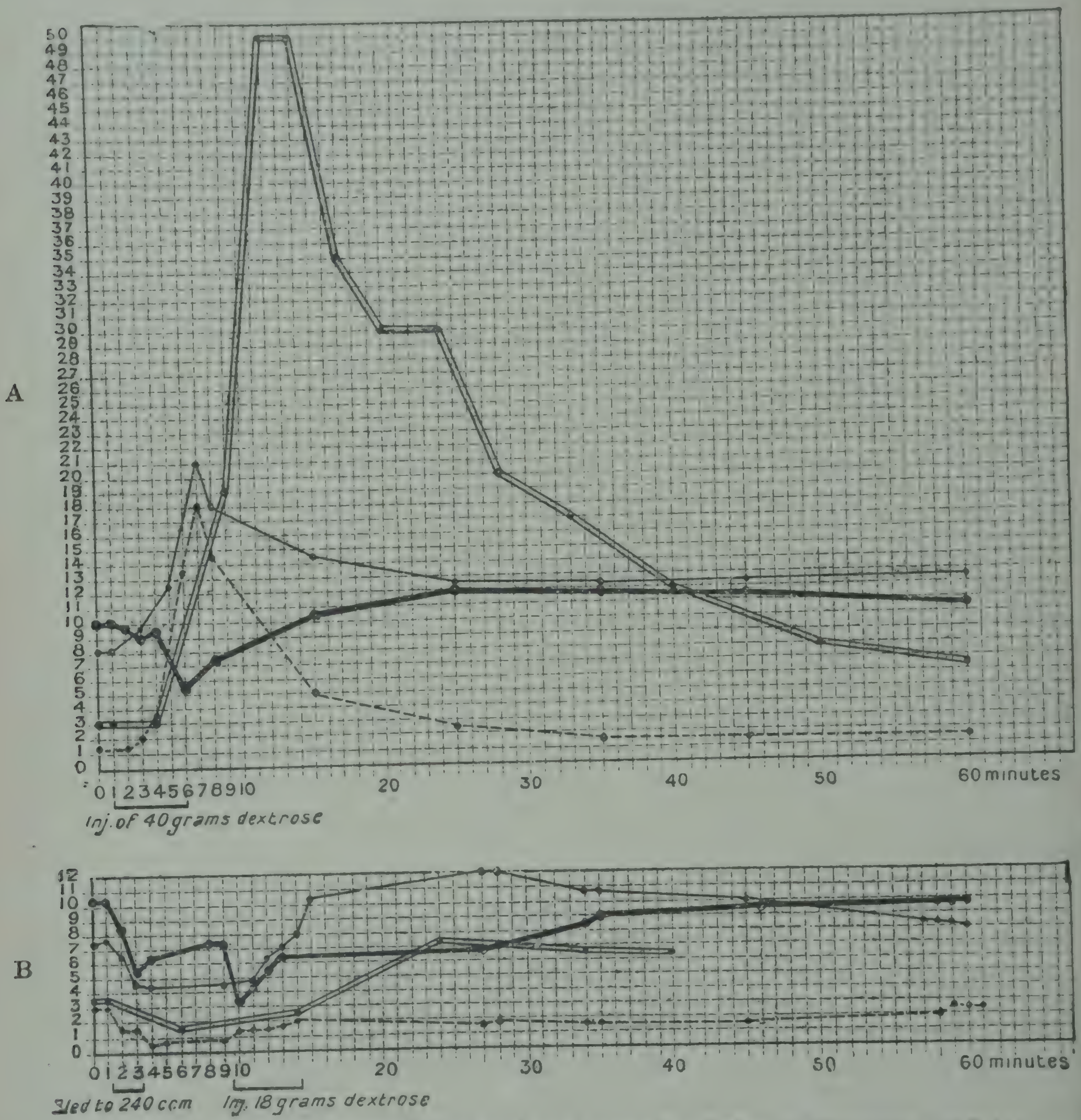


FIG. 470. Effect on Lymph Flow and on Arterial and Venous Pressures of Injection of Concentrated Solution of Glucose.

In B 240 c.c. blood was drawn from the animal before the injection. The double line = lymph flow in c.c. per ten minutes; thin line = portal vein; thick line = carotid artery; dotted line = inferior vena cava. (STARLING.)

Movement of Lymph. In the frog, the circulation of lymph is maintained by rhythmically contracting "lymph hearts." In the higher animals, the onward flow of lymph is effected partly by the pressure at which it is passed out into the lymph capillaries, and by pulsations of nearby arteries, but also to a large extent by the contractions of the skeletal muscles. In the smaller lymph radicles the pressure of lymph may attain 8 to 10 mm. of water. In the thoracic duct, at the point where it opens into the great veins of the neck, the pressure is obviously about the same as in these veins, that is to say, from — 4 to 0 mm. Hg, the negative pressure being occasioned by the aspiration of the thorax. This difference of pressure is sufficient to cause a certain amount of flow, though under normal circumstances no lymph at all flows from resting limbs. The only part of the body which gives a continuous stream of lymph during rest is the alimentary canal, the lymph of which passes into the lacteals, and thence makes its way through the thoracic duct. Movement, active or passive, of the limbs at once causes a flow of lymph from them.* Since the lymphatics are all provided with valves (Fig. 471), the effect of external pressure on them is to cause the lymph to flow in one direction only, viz. towards the thoracic duct and great veins. Hence we may look upon muscular exercise as the greatest factor in the circulation of lymph. The flow of lymph from the commencement of the thoracic duct in the abdominal cavity to the main part of it in the thoracic cavity is materially aided by the respiratory movements ; since, with every inspiration, the lacteals and abdominal part of the duct are subjected to a positive pressure, and the intrathoracic part of the duct to a negative pressure.

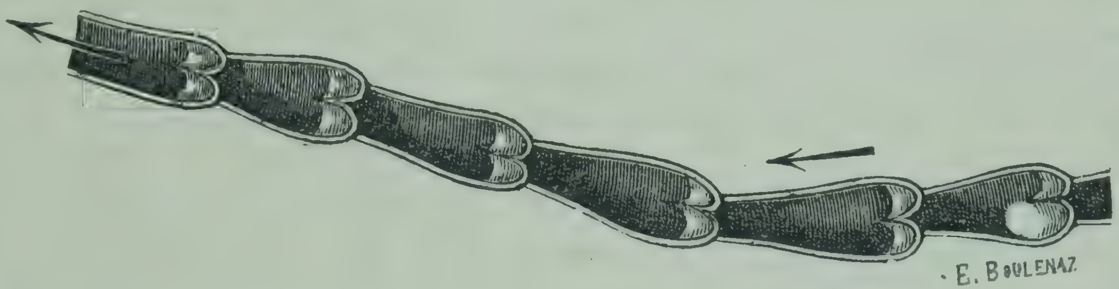


Fig. 471. A Lymphatic Vessel laid open to show Arrangement of the Valves. (Testut.)

Lymph Glands. Where a lymphatic gland is interpolated on the lymph path, a number of afferent lymph vessels enter it, and discharge the lymph into cortical lymph channels. After percolating through the gland the lymph is collected from the medullary part by a single efferent vessel ; the emergent lymph contains at least thirty times more lymphocytes than were present in the afferent stream. These latter few have escaped from the capillary blood.†

Lymphatic glands serve to collect particles and inactivate bacteria and toxins carried to them in the lymph stream (*e.g.* 99 per cent. of streptococci introduced into the lymph stream can be removed by passage through a lymph gland) : this is effected by the lymphocytes and monocytes in the gland, and it is often found that when tissues become infected with micro-organisms, the down-stream lymph glands become swollen and painful.

If a lymph gland is partially excised, the tissue regenerates, but after complete removal there is no re-formation of the gland. Owing to the ease with which new lymphatic vessels are formed, the removal of even a considerable number of lymphatic glands in an area is not usually followed by lymphatic obstruction.

Absorption from Lymph and Tissue Fluids.‡ On injecting certain coloured solutions (*e.g.* T.1824), into the connective tissues of a limb, and gently kneading the part, it is found that after a short time the fluid fills all the lymphatic channels running from the part, and we can in this way inject the lymphatics of the limb and trace their course on to the thoracic duct. If the limb is immobilised, the absorption does not occur. The same path is taken by micro-organisms as they spread in the body, even from fresh wounds, and by colloids and fine suspensions in general. Even the coarser particles of carmine or Indian ink which have been introduced in tattooing are slowly removed *viâ* the lymphatics, and deposited in lymph glands.

---

* White, Field and Drinker. *Amer. J. Physiol.*, 1933, **103**, 34.

† Drinker and Yoffey. *Anat. Rec.*, 1939, **74**, 417.

‡ Starling and Tubby. *J. Physiol.*, 1894, **16**, 140; Starling. *J. Physiol.*, 1896, **19**, 898, **22**, *Proc.*, xxiv.

This process of lymphatic absorption is a slow one, unless aided by passive or active movements of the surrounding parts, and cannot therefore account for the rapid symptoms of poisoning which supervene within two or three minutes after the hypodermic injection of a solution of a soluble poison. The absorption of crystalloids is into the blood stream and not into the lymph, as is shown by the fact that the symptoms occur almost as quickly when all the tissues of an injected limb have been severed, with the exception of the main artery and vein. In the absorption of fluids from the pleural and peritoneal cavities, movements, such as those of respiration, have been supposed to play a part by aspirating fluids into the lymphatic capillaries. Yet, after injecting methylene blue or indigo carmine into the pleural cavity or subcutaneous tissues, the dye-stuff appears in the urine long before any trace of colour can be perceived in the lymph flowing from the thoracic duct.

Whether a substance introduced subcutaneously is absorbed into the blood stream direct, or into the lymph, is largely a question of its molecular size.* Thus, Russell viper venom, diphtheria toxin and tetanus toxin (mol. wt. $>$ 20,000), and tiger snake venom (mol. wt., 20,000) pass into the lymph, as do many colloidal dyes, while cobra venom (mol. wt. $<$ 5,000) and crystalloids generally, pass into the blood stream.

Definite, though slow, passage of fluid, of colloidal dyes, of toxins and even of bacteria has been shown to occur from the nasal passages into the cervical lymph, especially if passive movements of the head are made.

More difficulty is presented by the question of the mechanism of absorption, into the blood vessels, of the normal tissue fluids—such an absorption as we have seen to occur after loss of blood by hæmorrhage. It seems probable that this absorption depends on the small proportion of protein contained in the tissue fluid as compared with the blood plasma, and is due to the osmotic pressure of the protein. Thus, we may conceive that there is normally a balance in the capillaries between the processes of exudation and of absorption, the former being conditioned by the capillary blood pressure and the latter by the difference in protein content, and therefore of colloidal osmotic pressure, between the blood plasma and tissue fluid. A rise of capillary pressure will upset this balance in favour of transudation, and the blood will become more concentrated; whereas a fall of pressure will turn the scale in favour of absorption, and the volume of blood will be increased at the expense of the tissue fluids.

Serous Fluids. The pericardial and peritoneal fluids are apparently ultrafiltrates of plasma as regards their electrolyte content. The following shows the kind of relationship :—

AVERAGE FOR SERUM, PERICARDIAL FLUID AND PERITONEAL FLUID OF DOGS. (MAURER, WARREN and DRINKER.†)

	Serum	Pericardial Fluid	Peritoneal Fluid
Protein per cent. . .	5·93	1·82	2·56
Chloride mg./100 c.c. .	412	440	448
Ratio $\frac{\text{Cl}' \text{ serum}}{\text{Cl}' \text{ fluid}}$ . .	—	0·94	0·93

The albumin/globulin ratio, generally higher than in blood, is variable. The fluids clot on removal.

---

* BARNES and TRUETA. *Lancet*, 1941, **240**, 623.

† MAURER, WARREN and DRINKER. *Amer. J. Physiol.*, 1940, **129**, 635.

For Reference

CHAMBERS and ZWEIFACH (1947). Intercellular Cement and Capillary Permeability. *Physiol. Rev.*, **27**, 436.

CUNNINGHAM (1926). The Physiology of the Serous Membranes. *Physiol. Rev.*, **6**, 242.

DRINKER and YOFFEY (1941). Lymphatics, Lymph and Lymphoid Tissue. (Harvard Univ. Press.)

DRINKER (1938). The Functional Significance of the Lymphatic System. *Harvey Lects.*, 1937–8, p. 89.

DRINKER (1942). The Lymphatic System. (Stanford Union Press ; Oxford Univ. Press.)

McLEAN (1925). Edema as a Problem in Physiological Regulation. *Physiol. Rev.*, **5**, 618.

McMASTER (1941). Lymphatic Participation in Cutaneous Phenomena. *Harvey Lects.*, 1941–42, p. 227.

STARLING (1909). The Fluids of the Body. (London : Constable.)

THE CEREBRO-SPINAL FLUID

Four membranes are interposed between the bony wall of the cranial cavity and the brain itself. From outside inwards, these are : (1) the periosteal layer of the dura ; (2) the meningeal layer of the dura ; (3) the arachnoid and (4) the pia. The cerebro-spinal fluid lies in the space between the arachnoid and the pia.

The dura consists of firm fibrous tissue ; its outer part represents the periosteum, and this is firmly attached to the interior of the cranium. Fibrous partitions are sent from the dura into the cavity of the cranium, thus supporting the chief parts of the brain. In the spinal canal, the bones have their own periosteum, and the meningeal dura, which is closely attached round the margins of the foramen magnum, forms a loose sheath round the spinal cord. Tubular prolongations are formed where each nerve-root emerges, and are continuous with the outer sheaths of the nerves. Where the venous sinuses occur, the two layers of the dura separate, so that the venous sinus lies between them. These sinuses are angular clefts, the chief of which lie along the attached margins of the falx cerebri and the tentorium cerebelli. The greater part of the blood which these contain drains into the internal jugular veins. In the spinal cord, these venous sinuses are replaced by plexuses of thin-walled veins, which are embedded in fat and lie on the outside of the dura. The arachnoid is a delicate, transparent membrane, covering the whole of the brain and spinal cord. Superficially it is covered by a layer of endothelial cells which bounds the sub-dural space. On its deep surface, it is connected with the pia mater by numerous fine fibres. It bridges over the inequalities in the surface of the brain, so that spaces are left which are filled with cerebro-spinal fluid. The arachnoid surrounds the cerebral and spinal nerves, and encloses them in loose sheaths as far as their points of exit from the skull or vertebral canal. On the upper surface of the brain it is very thin ; at the base it is thicker. The spinal part of the arachnoid is thin and delicate. It consists histologically of white fibrous and elastic tissue intimately blended, and contains blood vessels and a rich plexus of nerves. Wherever it comes in contact with cerebro-spinal fluid it is covered by flattened polygonal mesothelial cells.

The pia mater is a layer of connective tissue which serves to carry the blood supply to the whole surface of the brain. It closely follows all the irregularities of the latter, dipping down into the fissures and sulci of the brain. In the spinal canal, the pia sends out a series of processes, the ligamentum denticulatum. The outer extremities of these are attached to the dura mater, and help to sling the spinal cord in its dural sheath. By means of the rich blood supply conveyed by the pia mater, the brain is abundantly supplied with blood. From the circle of Willis, three main arteries on each side pass to supply separate regions of the brain.

There is a considerable space between the arachnoid and the pia covering the spinal cord, and since the cord ends at the first lumbar vertebra, whereas the dura and arachnoid extend to the second sacral vertebra, it is easy to puncture the theca and so to draw off cerebro-spinal fluid by means of a hollow needle. This simple operation, called *lumbar puncture,* is usually performed by piercing with the needle at the interval between the fourth and fifth lumbar vertebræ, and is of considerable value for diagnostic purposes as well as for the induction of " spinal anæsthesia," by the introduction of appropriate local anæsthetics into the thecal space.

THE SUB-ARACHNOID CISTERNÆ. The sub-arachnoid cavity, which is the space between the arachnoid and pia, contains the cerebro-spinal fluid and the larger blood vessels of the brain. It is traversed, as we have said, by spongy tissue consisting of delicate trabeculæ, which run from arachnoid to pia. In most parts of the cranial cavity, the space thus formed is small ; in certain parts near the base of the brain, however, the arachnoid is separated from the pia by wide intervals which communicate freely with one another and are named the sub-arachnoid cisternæ, *e.g.* the *cisterna magna,* below the cerebellum. The whole of these cavities, which contain the cerebro-spinal fluid, are lined by flattened mesothelial cells.

The Choroid Plexuses. These plexuses, which lie in the ventricles of the brain, consist of highly vascular tufts covered by ependymal cells. These peculiar masses of tissue apparently form a good deal of the cerebro-spinal fluid. Evidence of the secretory function of the choroid plexuses is partly of an histological, and partly of an experimental, nature. The cells are seen to contain granules, both eosinophil and basophil. Pilocarpine causes the protoplasm of the cell to become greatly increased in amount. The experimental evidence is that if, by blocking the foramen of Monro, one of the ventricles is prevented from getting rid of any fluid secreted into it, then hydrocephalus develops in that particular ventricle. Such a process does not take place if the choroid plexus be first removed from the ventricle.* Mechanical filtration no doubt contributes to the formation of C.S.F. Normally there is a constant flow of fluid from the lateral ventricles, so that it can be drawn off continuously by a catheter passed through the aqueduct of Sylvius into the third ventricle.†

It has been suggested that the rate of formation of cerebro-spinal fluid is accelerated by any circumstance which produces an increase of capillary pressure in the choroid plexuses. Yet occlusion of the great vein of Galen, which drains the choroid plexus, is not usually followed by hydrocephalus.‡ On the other hand, pressure may be exerted on the vein of Galen by reason of the vascular engorgement of a small body, the suprapineal arachnoid body,§ possibly under nervous control, and thereby the capillary pressure in the third and lateral ventricles, and the formation of C.S.F., may be increased.

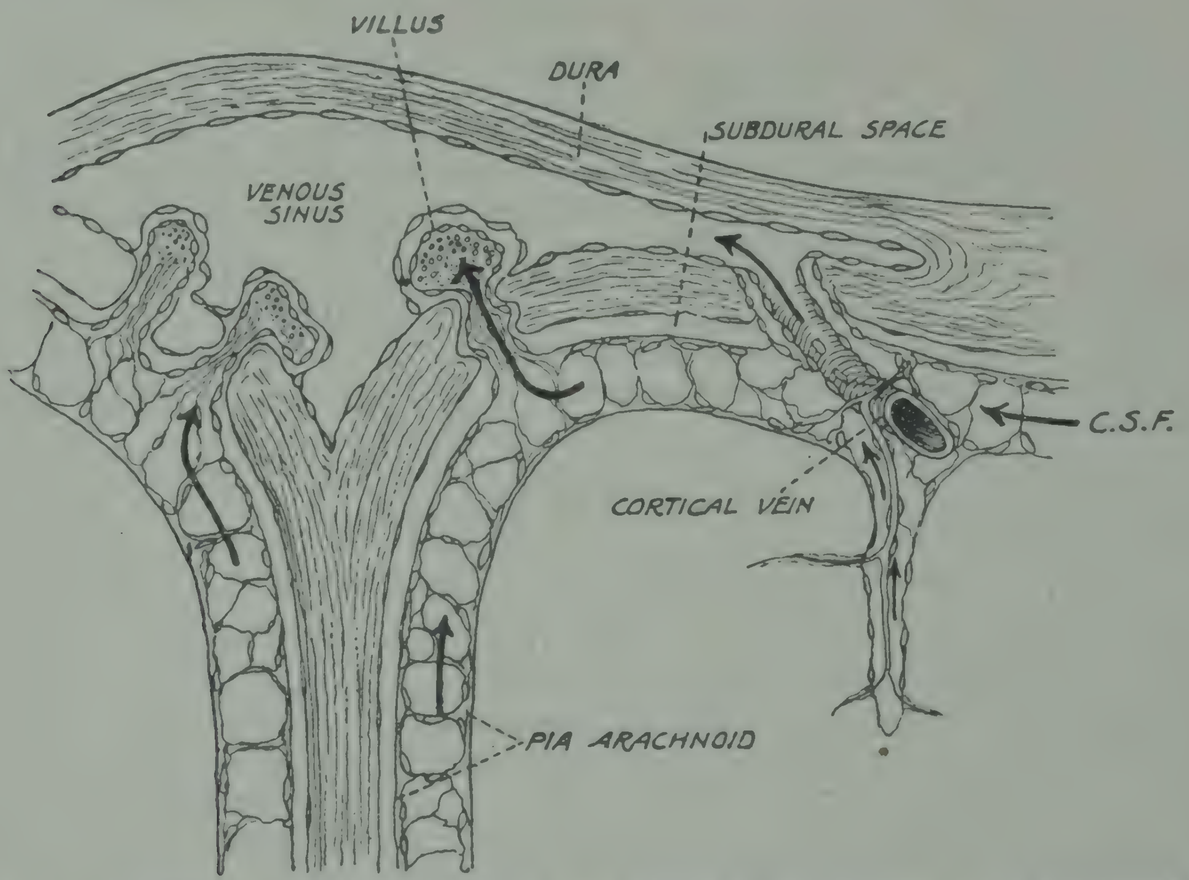


Fig. 472. Diagram of Relations of Sub-arachnoid Space to Arachnoid Villi, Cortical Veins, and Venous Sinus. (After Weed and Cushing.)

The Arachnoid Villi. These are minute processes of the arachnoid which project into the venous sinuses. Where these projections occur, the dura mater is interrupted. The processes themselves are lined by two layers of cells. Externally they are covered by the endothelium of the venous sinuses, and internally by the mesothelial cells of the arachnoid mater. They are present in most animals, as well as in men of all ages, are present in most of the cerebral sinuses, and are occasionally found penetrating the dura in other regions.

The Perivascular Spaces. Surrounding each of the blood vessels which penetrates into the surface of the central nervous system, is a sleeve formed of mesothelial cells. These tubes are called perivascular spaces. Thus a path is left clear for fluid to escape from the cell spaces inside the brain substance into the cerebro-spinal space. If these perivascular channels are followed inwards, they terminate in innumerable spaces, called perineuronal spaces, in which lie the nerve cells.

---

* Dandy. *Trans. Amer. Surg. Assoc.*, 1919, **37**, 397.

† Weed. *J. Med. Res.*, 1914, **26**, 93.

‡ Bedford. *Brain*, 1934, **57**, 1, 255.

§ Le Gros Clark. *B.M.J.*, 1940 (i), 476.

The relations of the perivascular spaces and arachnoid villi to the cerebro-spinal fluid and the blood are shown in Fig. 472.

The Cerebro-spinal Fluid. The chemical composition of this fluid is rather like that of lymph. Its specific gravity is about 1005. It contains salts, glucose and a little protein. Urea and creatinine are present in small amounts. Although at first sight the cerebro-spinal fluid might appear to be an ultra-filtrate or a dialysate of blood plasma, the balance of opinion seems to be that in its detailed composition it differs materially from those and must be regarded as a true secretion. Its total quantity in man is about 150 c.c. Its pressure inside the cranial cavity is about 100 mm. of water in recumbent subjects, but varies with posture ; usually it is definitely higher than the venous pressure in the torcula ; in the erect posture it may be below the atmospheric pressure. In disease, changes in the chemical composition and cells of the cerebro-spinal fluid may take place.

Circulation of Cerebro-spinal Fluid. This originates from two sources, the choroid plexuses in the ventricles, and the perivascular spaces found in relationship with the vessels, especially the arterial capillaries of the pia. It flows from the lateral ventricles by way of the foramen of Monro into the third ventricle, and thence, by the aqueduct of Sylvius, it reaches the fourth ventricle, from which it escapes, by way of the foramina of Majendie and Luschka, into the subarachnoid space. Having originated in this manner, the cerebro-spinal fluid passes slowly towards the convexity of the cerebrum to reach the arachnoid villi, through the walls of which it filters into the venous sinuses. That this is the normal direction taken by the circulation has been amply proved by Weed.* He injected into the lumbar subarachnoid space in animals a mixture of potassium ferrocyanide and ferric ammonium citrate, and, after an interval of several hours, fixed the entire brain in an acid fixative by which Prussian blue was precipitated wherever the mixture had travelled. The coloured precipitate was found all along the subarachnoid space and thence through the arachnoidal villi into the venous sinuses. There is no reason to suppose that the drainage is anything more than a filtration of the cerebro-spinal fluid from the subarachnoid space, where its pressure is higher, into the venous sinus, where its pressure is lower.

Under certain circumstances, it is possible to reverse the normal direction of flow. Instead of cerebro-spinal fluid being produced in the peri-vascular spaces, as is normally the case, in cerebral anæmia the spaces may be widely dilated and absorption of cerebro-spinal fluid may take place there, presumably into the capillaries and veins. Hence, if the ferrocyanide mixture be injected under these conditions, staining of the perivascular spaces, and even of the perineuronal spaces, results. A similar result follows after the intravenous injection of hypertonic salt solutions,† by which water is rapidly abstracted from the brain tissue as well as from the cerebro-spinal fluid, by way of the perivascular spaces. Owing to the rapid osmotic removal of water, the brain also shrinks appreciably, a fact which has been utilised in the performance of surgical operations on the brain, parts of which become much more readily accessible after a preliminary injection of strong salt solution. The total volume of cerebro-spinal fluid is augmented to keep pace with the shrinkage of the brain.

That there is some drainage of the cerebro-spinal fluid into lymphatics is suggested by the fact that if a colloidal dye is injected into the subdural space, dye very soon appears in the cervical lymph ducts.

---

* WEED. *J. Med. Res.*, 1914, **26**, 21, 51, 93.

† WEED and MCKIBBEN. *Amer. J. Physiol.*, 1919, **48**, 531.

For Reference

ASS. RES. NERV. MENT. DIS. (1924). The Human Cerebrospinal Fluid. (Baltimore: Williams and Wilkins Co.)

FLEXNER (1934). The Chemistry and Nature of the Cerebrospinal Fluid. *Physiol. Rev.*, **14,** 161.

FRIEDEMANN (1942). Blood-Brain Barrier. *Physiol. Rev.*, **22,** 125.

KATZENELBOGEN (1935). The Cerebrospinal Fluid and its Relation to the Blood. (London: Oxford University Press.)

WEED (1922). The Cerebrospinal Fluid. *Physiol. Rev.*, **2,** 171.

WEED (1933). Positional Adjustments of the Pressure of the Cerebrospinal Fluid. *Physiol. Rev.*, **13,** 80.

CHAPTER XXXVIII

THE DEFENCE OF THE ORGANISM AGAINST INFECTION

THE CELLULAR MECHANISMS OF DEFENCE

INVASION of a higher animal or plant by some lower organism is fraught with danger to the host. There are four types of lower organism against which defence is needed, viz. *viruses* (*e.g.* measles, small-pox), *bacteria* (*e.g.* tetanus, diphtheria), *moulds* (*e.g.* ringworm), and *protozoa* (*e.g.* malaria, amœbic dysentery).

The prevention of infection is of course the function of the epithelial covering, either of the skin, or of the surface of the gut. Protection here may be of a physical or chemical character. The cells may secrete a horny or chitinous layer, which presents a mechanical obstruction to the entry of bacteria. They may secrete mucin, which entangles invading micro-organisms, or they may secrete substances which actually kill such organisms. Thus, infection of the eyes is prevented by the presence, in the tears, of an enzyme, *lysozyme* * which is extremely potent in destroying most airborne organisms ; the same enzyme is also present in the mucus of the respiratory tracts, uterine cervix, etc., in skin, and indeed in most tissues.

When a micro-organism has obtained entrance through a wound, the task of dealing with the invader becomes the office of a special type of cell called a *phagocyte*, and the whole process by which particulate foreign material, or the animal's own dead tissues are got rid of, is spoken of as *phagocytosis*. It may be studied in its simplest form by injuring some tissue which is free from blood vessels. Thus, the tail fin of an embryonic axolotl may be cauterised with silver nitrate ; a certain number of cells are destroyed, and the dead tissue thereupon is treated as a foreign body. As a result, the wandering cells or histiocytes (macrophages) move from the surrounding tissues towards the seat of the injury, and the day after the injury has been inflicted, a collection of these macrophages can be seen, many of which contain débris of the destroyed tissue, which they have taken up. The cells finally wander away from the part, and the destruction is made good by the proliferation of the fibroblasts and of the epithelium immediately adjoining the injury.

Inflammation. It is evident that the accumulation of cells for the defence of the organism against invading microbes will be much more effective if the blood vessels participate in the process, so that, by their means, the phagocytic resources of all parts of the body can be drawn upon to ward off a localised attack. The process of phagocytosis thus, in the higher animals, becomes merged into the more complex series of phenomena to which the term "*inflammation*" has been applied. This process can be studied by observing the effects of slight injury to some vascular transparent part of the body, *e.g.* the frog's tongue or mesentery, or the web of the frog's foot. For this purpose a small piece of the skin of the frog's web may be snipped off with fine curved scissors, the section being sufficiently deep to remove the skin without causing hæmorrhage. The first effect noticed in the immediate neighbourhood of the injury is a dilatation of the vessels, especially of the venules, and an acceleration of the blood flow. In the course of an hour, the capillaries also become dilated, and many capillary channels, previously invisible, are now occupied with blood. In some instances, in higher animals, new capillaries have been observed, under the microscope, to grow out from other capillaries, towards the area of inflammation. Through the dilated capillaries there is a rapid blood stream, the corpuscles occupying the axis of the vessel, so that there is a periaxial layer of plasma. A little later, this acceleration gives place to a slowing of the blood stream, and simultaneously the leucocytes of the blood are seen to be adherent to the capillary wall. Apparently the latter becomes what we may call "sticky," the effect of the stickiness being to increase the resistance to the passage of the blood through the vessel, and also to cause the adhesion of the leucocytes to the wall. As the current becomes still slower, the distinction between axial and peripheral streams disappears. The corpuscles are now closely packed together, the white corpuscles being predominant at the margins of the capillary, where they form a lining to the vessel (Fig. 473). The next stage is the emigration of the leucocytes. These may be observed to thrust a process through the vessel wall (according to Arnold this process of emigration always occurs through the stigmata, *i.e.* the points where the endothelial cells come in contact—Fig. 474). The prolongation enlarges on the outer side of the vessel, while the portion of the leucocyte within the vessel becomes smaller, so that finally the whole leucocyte flows through, and lies in the tissue spaces outside the capillary.

---

* FLEMING. *Proc. Roy. Soc. Med.*, 1932, **26**, Path. Sec. 1.

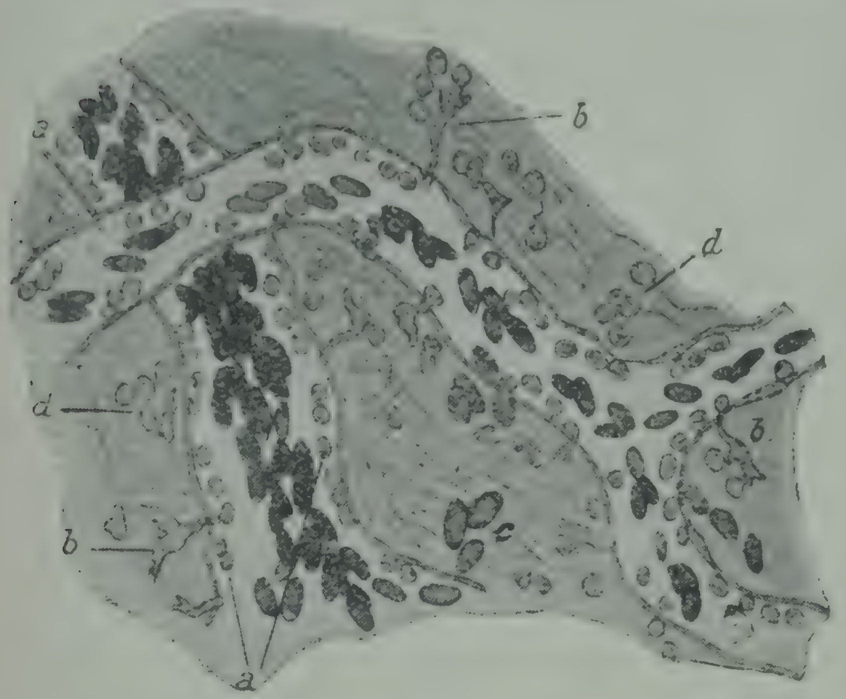


FIG. 473. Inflamed Mesentery of Frog, to show margination of leucocytes in the inflamed capillaries, *a*; migration of leucocytes, *b*; escape of red corpuscles, *c*; accumulation of leucocytes outside the capillaries, *d*.
(From ADAMI after RIBBERT.)

In the course of five or six hours, all the capillaries and small veins in the neighbourhood of the injury may show a crowd of polymorphonuclear leucocytes along their outer surfaces. In addition, there is a migration of certain tissue cells, viz. fibroblasts and histiocytes, to the affected place; later, there is often a further emigration from the blood of histiocytes and of lymphocytes. Of these various cells, the polymorphonuclear leucocytes and the histiocytes are phagocytic, and help to remove invading particles or organisms, and also later to remove the tissue injured by the primary lesion. As soon as this is effected, repair of the injured tissue occurs by a proliferation of the fibroblasts and the epithelium, while the leucocytes are destroyed by macrophages.

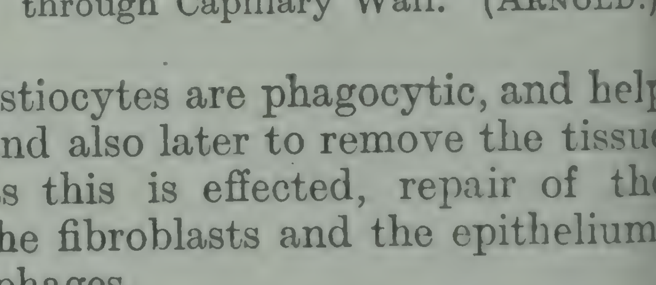


FIG. 474. Emigration of Leucocytes through Capillary Wall. (ARNOLD.)

Inflammation in warm-blooded animals always gives rise to dilatation of vessels and increased vascularity of the part, to alteration of the vessel wall and therefore to increased effusion of fluid. There are warmth and redness of the part from the vascular dilatation, swelling from the effusion of tissue fluid, and very often, as a result of the injury or the swelling and the consequent involvement of sensory nerves, pain. The four cardinal symptoms of inflammation, namely, *rubor*, *calor*, *turgor* and *dolor*, which have been described for generations as typical of this condition, naturally leave out of account the phenomenon of phagocytosis round the seat of injury with the objects of removing injured tissue, of destroying micro-organisms, of protecting the body from general infection, and of preparing the way for reintegration of tissue.

The essential phagocytic character of the inflammatory process is clearly shown if the primary lesion be attended with infection. Thus, if a small quantity of a culture of staphylococcus be injected into the subcutaneous tissue of the rabbit, the vessels surrounding the point of injection may, within four hours, be found densely filled with corpuscles. In ten hours' time the leucocytes are present in large numbers outside the vessels, while the injected cocci have spread for some distance along the lymphatic spaces and, while still partly free, have been to a large extent ingested by the leucocytes. In twenty hours' time the connective tissue fibrils at the point of injection are found to be widely separated by the aggregation of phagocytes. In forty-eight hours' time a well-defined abscess is produced. At the centre of this, all traces of previous connective tissue have disappeared and its place has been taken by *pus*, *i.e.* a mass of leucocytes, many of them dead and in a state of degeneration, mingled with staphylococci, partly within, partly outside the cells. The margin of the abscess is formed by connective tissue infiltrated with living macrophages. A certain number of cocci are to be seen free in the tissue outside this layer, but in the course of a day or two these free cocci disappear, and there is thus a continuous layer of phagocytes surrounding the abscess cavity, and preventing any further invasion of the body as a whole from the seat of infection. The abscess subsequently discharges on to the exterior by a process of necrosis of the overlying skin, and regeneration of tissue takes place in the same manner as after the more trivial injury.

We have abundant evidence to show that the essential factor in this aggregation of leucocytes is their chemical sensibility, and that the phenomenon is simply one of *chemotaxis*. A capillary glass tube containing a suspension of dead micrococci, or peptone, if introduced into the anterior chamber of the eye or into the subcutaneous tissue, is found after a short time to be full of leucocytes. The positive chemotactic influence exerted by any given species of pathogenic bacterium is roughly inversely proportional to its virulence. A culture lacking in virulence may cause a very pronounced aggregation of leucocytes which speedily destroy the micro-organism, whereas if a culture of a more virulent variety of the same microbe be injected, there may be all the signs of inflammation, swelling, and large effusion of fluid, though the tissues may contain very few leucocytes. Under the circumstances, the micro-organism rapidly proliferates and spreads from the seat of the lesion, giving rise finally to general infection. Usually, however, when bacteria enter the blood stream, they are rapidly removed by fixed histiocytes of liver and spleen, by polymorphonuclear cells in the blood, and by macrophages in the lungs.

The increase in capillary permeability and the chemotactic action on leucocytes appear to be due to the presence of a crystalloid substance formed in the tissues, and to which Menkin has given the name *leukotaxin;* it is probably a polypeptide.* It can be extracted from inflammatory exudate from which the cells have been removed. Such an exudate can be obtained from the peritoneal cavity after injection into it of turpentine. Injection of leukotaxin reproduces many of the phenomena of inflammation, *e.g.* increased permeability of capillaries, and migration of polymorphs, and, later, of macrophages. The local effect on capillaries can be strikingly shown by giving an intracutaneous injection on the rabbit's abdomen, followed by an intravenous injection of trypan blue. Around the site of leukotaxin injection the skin speedily turns blue, but remains unchanged round the site of a control injection of saline.

There is no doubt that, in all active inflammations, the polymorphonuclear leucocyte is the form which is attracted first, and in largest numbers, to the seat of injury. It is the characteristic cell from which pus is formed, and is actively phagocytic. It has nothing to do with the regeneration of the destroyed tissue. The eosinophile corpuscle is also present at an early stage around the inflammatory focus, but is never present in numbers at all comparable with those of the polymorphonuclear leucocyte. It is especially abundant in chronic inflammations of certain tissues, such as the skin, but is rarely seen to ingest bacteria, and therefore cannot be spoken of as phagocytic. The lymphocytes predominate in certain chronic inflammations, especially in those caused by the tubercle bacillus. They do not ingest bacteria. The histogenous wandering cells (macrophages) appear in the inflammatory area at a later period than the polymorphonuclear and eosinophile cells. They are actively phagocytic and motile. As a rule, their phagocytic properties are exerted, not on bacteria, but on other cells and cell *débris*. After an acute inflammation, their chief office is to clear away the remains of the polymorphonuclear leucocytes and dead tissues, so as to prepare the way for subsequent regeneration. The fibroblasts are connective tissue cells, and it is held by some investigators that they can be converted into macrophages, and *vice versâ*. It is, however, difficult to be certain whether the wandering and the fixed connective tissue corpuscles are of identical or of different origin. In the process of repair, the fibroblasts are converted into fibrous, or scar, tissue. We thus see that several types of the wandering cells of mesodermal origin, which take part in inflammation, do not destroy invading organisms by the process of ingestion. Yet the part played by such cells in the defence of the organism is no less important than that of the actively phagocytic cells.

---

* Cullumbine and Rydon. *Br. J. Exp. Path.*, 1946, 27, 33.

Leucocytes are able to destroy bacteria, not only by the process of intracellular digestion, but also by the excretion, into the surrounding body-fluids, of substances which have a deleterious influence on bacteria. Thus, normal blood serum is found to have a strong destructive influence on most species of bacteria, whether pathogenic or not. Since this property is not shared to anything like the same extent by the blood plasma, it may be ascribed to changes in the leucocytes in the process of clotting and the consequent liberation of bactericidal substances.

Many species of pathogenic bacteria cannot infect the animal as a whole. These, nevertheless, may multiply on the surface of the body, or in an abscess cavity, and lead to the death of the host, in consequence of the production by the bacteria of soluble *toxins*, which are absorbed into the blood stream. Examples of such micro-organisms are those which are associated with tetanus and diphtheria. There must be other mechanisms at the disposal of the body for the neutralisation of these toxins. The protection of the body against destruction by bacterial toxins involves in fact a whole series of chemical mechanisms, of equal importance with the phagocytic mechanism.

THE CHEMICAL MECHANISMS OF DEFENCE

Infection and Immunity. All infectious diseases are caused by the agency of micro-organisms. The greater number of these are bacteria or viruses, a certain number must be classed with the yeasts, while others are protozoal in character.

In the protozoal diseases, *e.g.*, malaria, the micro-organisms exist for the greater part as intracellular parasites. One attack of the disease does not as a rule confer immunity, and the treatment has to be sought along the lines of medication by drugs, which will either directly destroy the infecting agent, or enable the tissue to combat it.

The infections due to the viruses are in many ways similar to those induced by bacteria, but in this case the causative " organism " is a substance which can pass through the pores of the finest filter, and is invisible under the microscope. Viruses are proteins of enzyme-like properties ; some have been crystallised, and shown to possess spherical or rod-like molecules, visible by electron microscopy : they cannot be cultivated except in living tissues,[*] whereas the bacterial micro-organisms can usually be cultivated on artificia media outside the body.

---

[*] STEINHARD and LAMBERT. *J. Infect. Dis.*, 1914, **14**, 87 ; MAITLAND and MAITLAND *Lancet*, 1928 ii., 596.

It is especially in the class of diseases due to bacteria, that the organism has developed chemical mechanisms of defence. The diseases due to bacteria include diphtheria, tetanus, tubercle, anthrax, typhoid and many others. In these diseases we have to deal with a number of phenomena more or less common to all. The infection in each case is due to the actual transference of the specific organism from one animal to another. After the micro-organism has attained entrance into the system there is a period of incubation before the disease is manifest. When this occurs, the specific microbe is to be found in large quantities either in the blood or in the tissues of the body. The disease is generally characterised by fever and often by local lesions, such as the intestinal ulcers of typhoid, or the glandular swellings of bubonic plague. The micro-organisms may develop in the animal until its death, or the disease may terminate in recovery and the total disappearance of the microbes from the body. Some animals or individuals are naturally immune to particular diseases, *e.g.* dogs to typhoid, and this is called *natural congenital immunity*. In other cases, it is found that, after recovery, the patient cannot be reinfected by the bacterium which was the cause of the disease, and this condition, called *natural acquired immunity*, may last as long as the patient lives.

Toxins. One class, of which the diphtheria and tetanus bacilli are examples, secrete into the surrounding culture-fluid substances, which, when injected into animals, act as virulent poisons. The other class of bacteria do not form such extracellular toxins, but in their case it is found that, if the bodies of the bacilli be broken up, the injection of the contents of the bacteria is attended with poisonous effects. The bacteria may be thus classified according as they produce *extracellular* or *intracellular toxins.*

We may first deal with the manner in which the body reacts to the exotoxins. If a culture of diphtheria or tetanus bacilli be filtered, the clear filtrate free from bacilli is found to exercise as poisonous results as if the culture of the living bacilli had itself been employed. The toxins contained in these fluids are extremely potent. Thus five-millionths of a gramme of dried tetanus toxin is a fatal dose for a mouse, and 0·00023 g. would kill a man; the pure toxins, which have never been isolated, must be much more powerful. The toxin is an unstable body, and is destroyed by heating to 65° C. The nature of toxins is not known, though they are probably proteins. Similar toxins are widely distributed. Thus they form the active constituent of snake venom and of the poison of scorpions and spiders. They also occur in the seeds of castor oil and of jequirity, the toxins of which are proteins, known respectively as *ricin* and *abrin.* There is a great variability in the reaction of different animals to these toxins. To the poison of tetanus, the rabbit is, weight for weight, two thousand times, and the hen twenty thousand times, more resistant than the guinea-pig. A certain time is necessary after the introduction of the toxin before its effects are displayed. There is a striking difference in this respect between the action of these complex bodies and the action of crystalloid drugs, such as strychnine. This time cannot be reduced beyond a certain limit, however much toxin be injected. Thus, a lethal dose of diphtheria toxin kills a guinea-pig in fifteen hours. If ninety thousand lethal doses be injected into a guinea-pig, it is not possible to reduce the time of survival below twelve hours. Another characteristic of these toxins is the specificity of their action. One kind of toxin may act chiefly on the central nervous system, another on the peripheral nerves, another on the red blood corpuscles. Associated with this specific action is the actual combination which occurs between the toxin and the organ on which it exerts its effect. Thus, tetanus toxin has a specific affinity for the central nervous system, and may be removed from a solution by shaking the latter with an emulsion of the brain of a susceptible animal, such as a guinea-pig, but not with one of the brain of an animal of a non-susceptible species.

Antitoxins. In spite of the excessively lethal character of these toxins, it is possible to render an animal immune to their action. If a dose of diphtheria or tetanus toxin which is smaller than the fatal dose be injected into an animal, the latter shows signs of injury, from which it recovers. When recovery is complete, it is found that three or four times the fatal dose may be injected without producing any evil effects ; and this process of injection of toxin may be repeated in continually increasing doses, until the animal may finally be able to withstand a dose one hundred thousand times as large as that which would have been fatal to it in the first instance. This is a condition called *active artificial immunity*, and it is found that the blood serum of the immunised animal has the power of neutralising the toxin. Thus, if blood serum from a horse which has been treated with large doses of diphtheria toxin be mixed with an equal quantity of the toxin itself, the mixture may be injected into susceptible animals without the production of any effect. It is possible in this way to get a serum, 1 c.c. of which will neutralise many fatal doses of the toxin ; further, this antitoxic serum may be injected into a susceptible animal and then confers *passive artificial immunity* on the latter, or it may be injected into a diseased animal and so used as a curative agent. *Antitoxins* thus play a great part in therapeutics, especially of tetanus and diphtheria. In the case of tetanus, the toxin appears to be absorbed by the motor end plates, from which it diffuses along the motor axons to reach the ventral horn cells of the cord. By the time that it has arrived at the central nervous system, and the spasms typical of tetanus have broken out, the toxin is already so firmly bound to the reacting tissue that the injection of antitoxin into the blood stream has little or no effect on the course of the disorder. The use of the tetanus antitoxin is, therefore, chiefly as a prophylactic agent.

If we define a unit of toxin as that amount which possesses a certain power, *e.g.* which will kill a guinea-pig in so many days, or will cause the complete hæmolysis of 1 c.c. of blood in two and a half hours, we can find the amount of anti-body which is just sufficient to neutralise this effect, and this amount of anti-body can be regarded also as one unit. If, instead of one unit of each, we take 100 units, the neutralisation is effected in the same way. The process is found, however, to be more complex when we take 100 units of toxin, or lysin, and attempt to neutralise them by the fractional addition of antitoxin. In the case of a strong acid and strong alkali, we know that if 100 c.c. of alkali are just sufficient to neutralise 100 c.c. of acid, the addition of 50 c.c. of alkali will leave half the acid unneutralised. If we try the same experiment in the case of mixtures of toxin and antitoxin, it will be found that the addition of 50 units of antitoxin will neutralise much more than half of the toxin, and the same applies to other bodies of this class. Further, 1 c.c. of antilysin exactly neutralises 1 c.c. of lysin ; but these two substances will no longer be in equilibrium when the whole is diluted up to 10 c.c. with water.

The interaction, therefore, resembles adsorption. The exact adsorption which takes place must be a function of the chemical configuration of the substance forming the surface, since otherwise it would be impossible to account for the extremely specific character of the interaction between toxins and their corresponding antitoxins. It probably depends upon the operation of such forces as electronic van der Waal's attraction, Coulomb attraction, electrical dipoles or multipoles, formation of hydrogen bonds, and other forces dealt with by physical chemistry. In any case, the antitoxin does not *destroy* the toxin, since it is possible, from an inert mixture of toxin and antitoxin, to separate toxin and anti-body once more.

THE SIDE-CHAIN THEORY. How are we to account for the production, as a result of the injection of toxins into the body, of antitoxin far transcending the amount of toxin injected ? In all the speculations on the mode of production and action of antitoxins an important part has been played by a conception put forward by Ehrlich in 1885. According to this conception, which is spoken of as the " side-chain theory," each unit of living matter consists of a centrally placed portion with a number of side chains attached to it. To these side chains, or *receptors*, various nutritive or other substances may be linked, provided their chemical configuration is of the right kind. Ehrlich regarded the toxins as possessing two side chains, one of which, by its stereomeric configuration, is peculiarly adapted to fit on to the side chains of the protoplasm which the toxin attacks, and this group is known as the *haptophore* group ; and another side chain, the *toxophore* group, which is responsible, when the toxin is once anchored, for the destructive changes wrought by the toxin on the cell of the body (Fig. 475, 2). The antitoxins are supposed to act by uniting with the haptophore group, and so preventing the susceptible protoplasm from doing so and suffering their injurious effects.

The formation of antitoxins is accounted for on this hypothesis in the following manner When a receptor of the cell is occupied by becoming attached to the haptophore group o the toxin, this side chain, is, so to speak, shut out from the normal activities of the cell. A defect is thus produced in the cell, to which the latter endeavours to adapt itself by the production of other side chains of the same character (Fig. 475, 3, 4). It may be regarded as a general rule in living tissues, that a reaction tends to be an over-reaction, so that the compensation by the cell should more than make good the defect produced by the attachment of the toxin. We thus get, not one, but a number of side chains produced, of the same character as that occupied by the toxin molecule, and therefore able also to act as receptors for the haptophore group of the toxin. These new receptor side chains, being produced in excess, were supposed by Ehrlich to be thrown off from the cell and to circulate in the body fluids (Fig. 475, 4). The fragments thus set free have a specific power of uniting with the toxin, and it is this excess of side chains, thrown off from the cell, which represents the antitoxin molecules found circulating in the blood after the injection of toxins. It will be noted that this theory, though chemical in form, is really purely biological. It does not explain the phenomena by reference to the known laws of chemistry, but is a manner of re-naming the biological phenomena, which facilitates their description and discussion and enables us to classify the very complex phenomena of immunity, though only in an imperfect fashion.

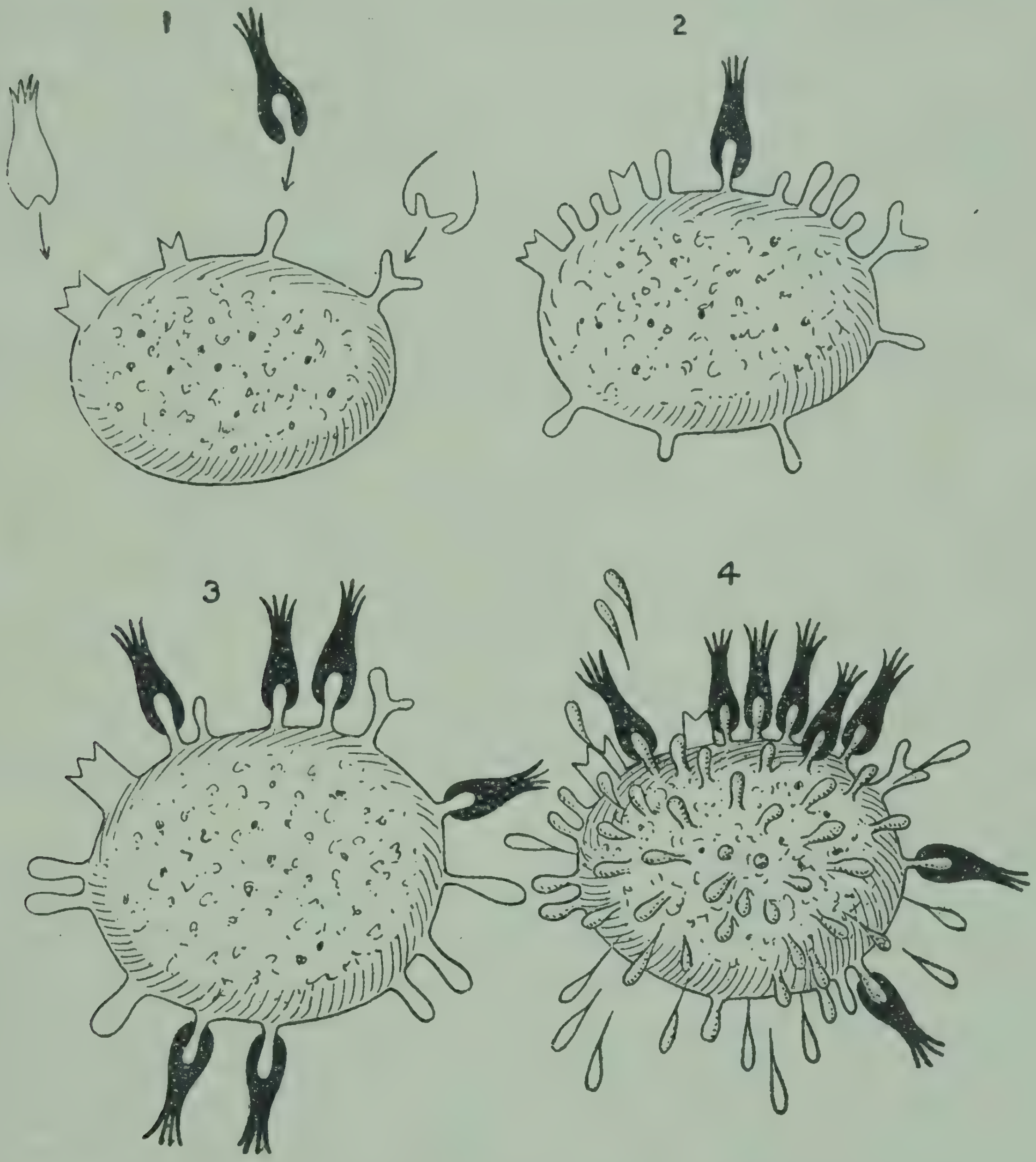


FIG. 475. Schematic Representation of Formation of Antitoxin as Side Chains of Protoplasmic Molecule. The black bodies are the toxin molecules which fit by their haptophore end on to the side chains of the cell. (EHRLICH.)

A more modern version of the hypothesis is that the antibodies are not formed in all cells of the body, but that the reticulo-endothelial cells, which are known to be active in taking up various antigenic materials, react by producing antibodies.

Antigens. The property of giving rise to anti-bodies on injection into an animal is not confined to toxins ; a large number of substances, all of them containing proteins, have the same property. All such substances are classed together as *antigens*, and are highly specific. Some proteins, *e.g.* insulin, gelatin, cannot act as antigens. Bacteria, alive or dead (and thus including endotoxins), or any animal or vegetable cells, venoms, viruses, or pure proteins derived from any of them, can act as antigens, provided they are foreign to the species of animal into which they are injected. (An exception is the protein of the crystalline lens, which can act as an antigen even for the same species.) They all give rise to the production of *specific anti-bodies*. Thus, human serum injected into a rabbit produces in a few days in the rabbit's serum a substance which will give a precipitate when mixed with human serum even in minute traces. This *precipitin* formation is specific, so that it may be used as a test to find the species of any unknown specimen of serum.

The antigenic properties of proteins may be artificially altered by the introduction of new groups: *e.g.* insulin and gelatin become antigenic when coupled with glucosidotyrosine; the resulting antibodies, however, do not react with the same antigen, but with the globulin analogue, glucosidotyrosyl-globulin.*

In a similar way, using an extract of dead bacteria, or a protein, as antigen, a specific precipitin for the bacterial extract or for the particular protein, as the case may be, develops in the blood.

A *vaccine* is such a collection of killed bacteria or virus, and is often injected into patients to stimulate the production of anti-bodies by their own tissues.

When live foreign cells, *e.g.* bacteria or red cells, have been injected, there is developed an anti-body which causes these cells, when added to the serum, to gather together or agglutinate. Typhoid fever can be diagnosed by the fact that the patient's serum causes agglutination of the bacteria from a pure typhoid culture (Widal test). The anti-body concerned is called an *agglutinin*.

The amount of an antibody formed after a single injection of an antigen varies with the quantity of antigen given, though not in direct proportion. After the injection, the antibody begins to appear in a few days, reaches a maximum about the fifteenth day, and then declines, with diminishing speed, over a period of months or years. Second doses of antigen cause earlier and larger antibody formation.

The nature of these specific anti-bodies, and the cause of their high specificity to the antigens which give rise to them, are not exactly known. They appear to be specialised serum γ-globulins, possessing specific groups which enable them to unite with the corresponding antigens. With regard to the antigens of bacteria, recent work has indicated that, while some of the specificity is attributable to proteins, the most specific immunological reactions are often associated with the presence of highly specialised polysaccharides.† These are, in fact, the "haptenes" postulated by Ehrlich. No antigenic properties are possessed by the specific polysaccharides when pure; to show such properties, a protein foreign to the reacting animal (*e.g.* one from the bacterial substance) must also be present.

Cytolysins. Normal goat's serum may be mixed with the red blood corpuscles of the sheep without any injury to the latter. If, however, washed sheep's erythrocytes be injected at intervals of a few days into a goat, the goat's serum is found to have acquired the power of rapidly agglutinating, and then of dissolving, the red blood corpuscles. A specific *hæmolysin* is said to have been produced. This hæmolytic power can be tested by mixing the serum and the washed blood corpuscles together and allowing the mixture to stand in a narrow tube. The corpuscles rapidly sink to the bottom, leaving the colourless serum above, unless hæmolysis has occurred, in which case the serum will be red.

If the hæmolytic serum be heated to 60° C., it is found to have lost its power of dissolving sheep's corpuscles. This power is at once restored if to the heated serum be added any normal blood serum, even of the sheep itself. It seems, therefore, that two substances are involved in the hæmolysis, namely, (*a*) a substance present in most normal sera, and believed to be formed by the leucocytes; it is destroyed at a temperature of 60° C. and has been called the *complement;* and (*b*) a substance which is only present in the serum as a result of the previous injection of some species of red blood corpuscle, which is resistant to the action of heat, and is called the *amboceptor*.

Apparently the amboceptor becomes attached to the red cell first, and then the complement is linked on to this and hæmolysis ensues. Thus, if a rabbit is sensitised against human red cells, and some of its serum taken and heated to 60° C., it will not now hæmolyse human red cells, because the complement has been destroyed by heat. If the human cells added to it are now removed by centrifugalisation, the addition of complement to them,

---

* CLUTTON, HARINGTON and YUILL. *Bioch. J.*, 1938, **32**, 1111.

† THOMSEN. "Antigens in the Light of Recent Investigations," 1931. (Copenhagen, Levin and Munksgaard.)

e.g. normal guinea-pig serum, causes immediate hæmolysis, whereas if complement is added to the supernatant serum removed from the red cells, no hæmolysis occurs on adding fresh human cells, because the amboceptor has all been removed by those first added.

The function of the amboceptor thus seems to be to enable the complement already present in normal serum to act upon the red blood corpuscles. We may regard the amboceptor, therefore, as having two haptophore groups, one of which anchors on to the red blood corpuscle, while the other attaches itself to the complement (Fig. 476, 7). The amboceptor *plus* the complement thus comes to resemble the toxin molecule, having a free haptophore group at one end and a toxophore group (the complement) at the other end. The reaction to the injection of the red blood corpuscles consists in the formation of the amboceptor, which is essentially the anti-body of the red blood corpuscle (Fig. 476, 8).

Similar specific anti-bodies, capable of effecting the dissolution of cells or organisms, may be produced by the injection of various species of bacteria, or of animal cells, such as leucocytes, spermatozoa, liver cells, etc., and there can be no doubt that bacteriolytic substances play a considerable part in acquired immunity.

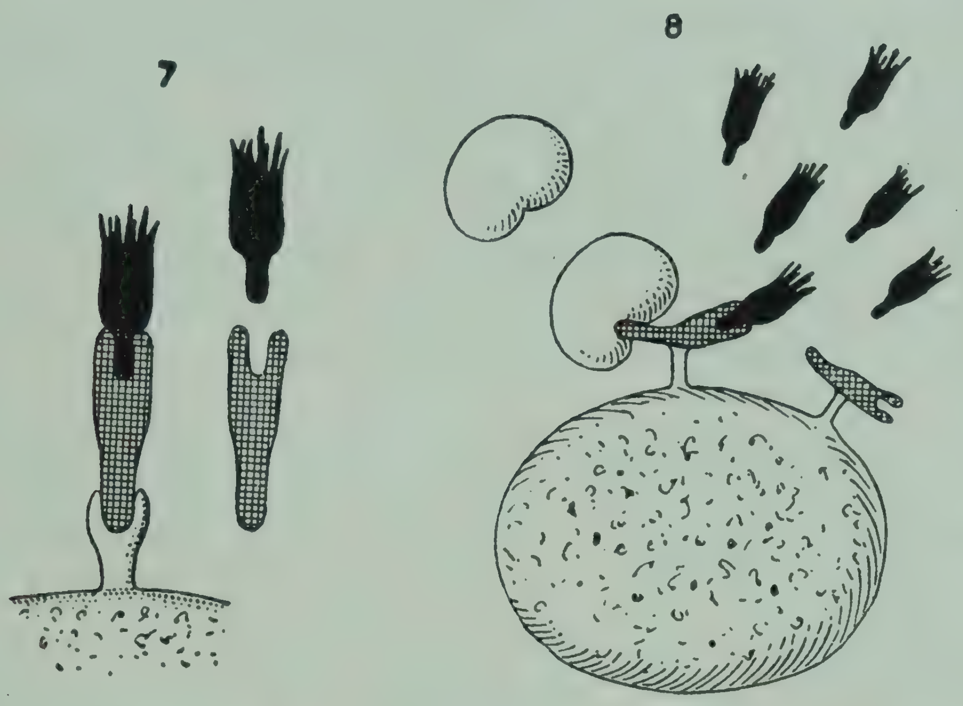


FIG. 476. Diagrams to show the relation of Amboceptor and Complement to Red Corpuscles (7) and to the Animal Cell (8). (EHRLICH.)

A mixture of red cells and specific hæmolytic amboceptor (heated immune serum) can be used as a test for the presence or absence of complement. If complement is added, hæmolysis occurs, as shown above. Based on this fact, and on the further fact that complement is fixed by the amboceptor in presence of its specific antigen, is the important *complement-fixation test* for the identification of either antigens or amboceptors. An important test related to this is the *Wassermann reaction* for syphilis; the patient's serum (= amboceptors ?) is heated to 55° C., mixed with certain lipoids (= antigens ?), and then some guinea-pig serum (complement) added. After a time, heated anti-sheep serum from a rabbit, and washed sheep corpuscles, are added, to test whether complement has been fixed or not. Hæmolysis indicates no fixation, and therefore a negative result of the test; absence of hæmolysis, a positive result.

Opsonins. In some cases the anti-bodies produced by the injection of living or dead micro-organisms do not bring about actual destruction of the bacteria, but alter them in such a way as to make them more susceptible to the action of the phagocytes. If washed white blood corpuscles be mixed with cocci, such as those found in an ordinary boil, they are found to take up the micro-organisms in considerable numbers. The numbers taken up are much increased in the presence of serum derived from an individual who has received repeated vaccine injections of the dead micrococci in question. To the substances in the serum, which thus prepare the micrococci for ingestion by the phagocytes, Wright gave the name of opsonins. The opsonic index of the leucocytes of any individual, in reference to a given species of microbe, is determined by observing the number of the microbes taken up by the leucocytes after treatment with the serum of the individual, and then comparing it with the number taken up by the same leucocytes when the bacteria have been treated with the serum of an average individual.

We thus see that immunity, whether innate or acquired, is extremely complex in character and may depend on one or more of many factors. The immunity of an animal to any given infection may be determined by the absence from his body of receptor groups for the toxin excreted by the microbe responsible for the infection, or by the fact that the receptor groups, though present, are confined to tissues on which the toxophore group can have no influence. Thus, *e.g.* an attachment of the tetanus toxin to a connective tissue cell would be without effect on the health of the body. Again, immunity may be due to the efficacy of the phagocytes, either of the fluids or of the connective tissues, in ingesting and destroying the micro-organisms, and this, as we have seen, may again be dependent on the presence in, or absence from, the body fluids of substances which, while not destroying the micro-organisms, render them more accessible to the action of the phagocytes. In those cases where the infecting organism secretes a specific toxin, the main line of defence, and the main factor in the production of immunity, is the formation of specific antitoxins to the poison in question. Finally, there may be produced, as a result of the excess of micro-organisms, substances such as the amboceptors, which render the micro-organisms susceptible to destruction by the complements or cytases normally present in the circulating fluids, and possibly themselves derived from the activity or destruction of the leucocytes and other phagocytes of the body.

In this short description we have been able to touch only upon the most salient features of the immunity problem. The question enters strictly into physiology, since, as we have seen, it involves adaptations on the part of the organism to changes in itself or in its environment.

Anaphylaxis (*v.* also p. 51). If an animal be given injections of a foreign protein, it is found that, after ten or more days, a small dose of the same antigen causes a very severe reaction, or death. This, which appears like the opposite to immunity, is called anaphylaxis. Death occurs owing to profound shock, which manifests itself differently in different species; thus in the guinea-pig death is due to asphyxia from constriction of the bronchioles. In the dog, death is due to engorgement of the liver from constriction of branches of the hepatic veins. These differences are explained by the differences in the abundance and distribution of plain muscle in different species, for in all cases generalised contraction of plain muscle is the outstanding feature of anaphylactic shock; Dale has shown that the plain muscle of a sensitised animal gives a powerful contraction when a trace of the antigen is added, even if the tissue has been previously washed free of all blood. The reaction is, therefore, due to a direct action of the antigen on the tissues, and there is much evidence in favour of the belief that the substances passed out as "precipitins" into the plasma are at a later stage absorbed by the tissue cells, and so sensitise these to the antigen. Thus, the anaphylactic sensibility is found to wax as the precipitin content of the serum wanes. In anaphylactic shock the blood coagulability is reduced or abolished; this is due to the liberation of heparin from the mast cells of the liver.[*]

The close similarity between the anaphylactic responses of different tissues in various species of animals, and the responses of those particular tissues to histamine, led to the suggestion that the phenomena of anaphylaxis were due to the liberation of histamine into the blood, and its subsequent distribution all over the body. The fact that histamine can be shown to be produced in the isolated tissues of sensitised guinea-pigs [†] gives support to the idea, but the anaphylactic response of isolated tissues discovered by Dale indicates that the histamine is locally produced. Lewis observed that the local anaphylactic response in the human skin was similar to the response to other skin injuries, and it now seems probable that the phenomena are due to the liberation, within the cells, of a histamine-like substance, in consequence of the injury produced by the union of anti-body and antigen.

Anaphylactic shock may be met with in medical practice, as when a second dose of a serum is given after a long interval. The name has also been loosely used to describe many different shock-like conditions, some of which are probably due to a direct toxic action on the vascular system, and in which there are no phenomena of antigenic sensitisation. The anaphylactic reaction is highly specific, and the antigens are invariably proteins. Benadryl, $(C_6H_5)_2CH—O—CH_2—CH_2—N(CH_3)_2$ (HCl)., competes with histamine, and is used as a drug to combat anaphylactic and allergic phenomenon.

For Reference

Arey (1936). Wound Healing. *Physiol. Rev.*, **16**, 327.

Baetjer (1932). The Effect of Muscular Fatigue upon Resistance. *Physiol. Rev.*, **12**, 453

Cannon (1940). The Functional Significance of Specific Agglutinins and Precipitins *Physiol. Rev.*, **20**, 89.

Dragstedt (1941). Anaphylaxis. *Physiol. Rev.*, **21**, 563.

[*] Jaques and Waters. *J. Physiol.*, 1941, **99**, 454.

[†] Schild. *J. Physiol.*, 1939, **95**, 393.

GAY (1924). Local Immunity. *Physiol. Rev.*, **4,** 191.
GUNN (1923). Cellular Immunity. *Physiol. Rev.*, **3,** 41.
HEIDELBERGER (1927). The Chemical Nature of Immune Substances. *Physiol. Rev.*, **7,** 107.
JAFFÉ (1931). The Reticulo-Endothelial System in Immunity. *Physiol. Rev.*, **11,** 277.
LONGCOPE (1923). Anti-Anaphylaxis and Desensitisation. *Physiol. Rev.*, **3,** 240.
McCUTCHEON (1946). Chemotaxis in Leukocytes. *Physiol. Rev.*, **26,** 319.
MARRACK (1938). The Chemistry of Antigens and Antibodies. Ref. No. 230, M.R.C. (London : H.M. Stationery Office.)
MENKIN (1938). Role of Inflammation in Immunity. *Physiol. Rev.*, **18,** 366.
MENKIN (1940). Dynamics of Inflammation. (New York : Macmillan Co.)
MUDD, McCUTCHEON and LUCKÉ (1934). Phagocytosis. *Physiol. Rev.*, **14,** 210.
NORDMANN (1931). Local Reactions in Sensitized Animals, etc. *Physiol. Rev.*, **11,** 41.
PAULING, CAMPBELL and PRESSMAN (1943). The Nature of the Forces between Antigen and Antibody and of the Precipitation Reaction. *Physiol. Rev.*, **23,** 203.
RIVERS (1932). The Nature of Viruses. *Physiol. Rev.*, **12,** 423.
SIMON (1923). The Filterable Viruses. *Physiol. Rev.*, **3,** 483.
STANLEY (1939). The Architecture of Viruses. *Physiol. Rev.*, **19,** 524.

CHAPTER XXXIX

RESPIRATION

I. THE MECHANICS OF THE RESPIRATORY MOVEMENTS

In multicellular animals tissue respiration is carried out between the internal medium, *e.g.* blood, and the individual cells ; and the use of the special organ of respiration is to bring the circulating internal medium in intimate relation, over a large area, with the surrounding air or water. In these animals we can distinguish two processes—viz. (1) the interchange of gases between the tissue cells and the surrounding fluid, " internal respiration " ; (2) the interchange of gases between the circulating fluid and the external medium, " external respiration." In most water animals the respiratory organ, the gills, presents a large surface, well supplied with blood, over which a continual stream of water is kept up.

In all air-breathing vertebrates the organs of external respiration, the lungs, contain numerous air sacs. The renewal of the air in the air sacs is effected by the movements of respiration, while a rapid circulation of blood is carried out through a fine meshwork of capillaries just underneath the surface of the sacs.

The lungs in their development grow out from the fore part of the alimentary canal into the front part of the pleural cavity. The surrounding body walls become strengthened by the formation of the ribs, so that the lungs are suspended in a bony cage-work, the thorax. Their outer surface is covered with a serous membrane, the *pleura*, which is reflected on to the wall of the thorax from the roots of the lungs, and completely lines the cavity in which they lie. The surface of the pleura facing the pleural cavity is lined with a continuous layer of flattened mesothelial cells, and is kept moist by the lymph-like pleural fluid. Thus, being attached to the thorax only where the bronchi and great vessels enter, the lungs are able to glide easily over the inner surface of the thorax, with which under normal circumstances they are in intimate contact.

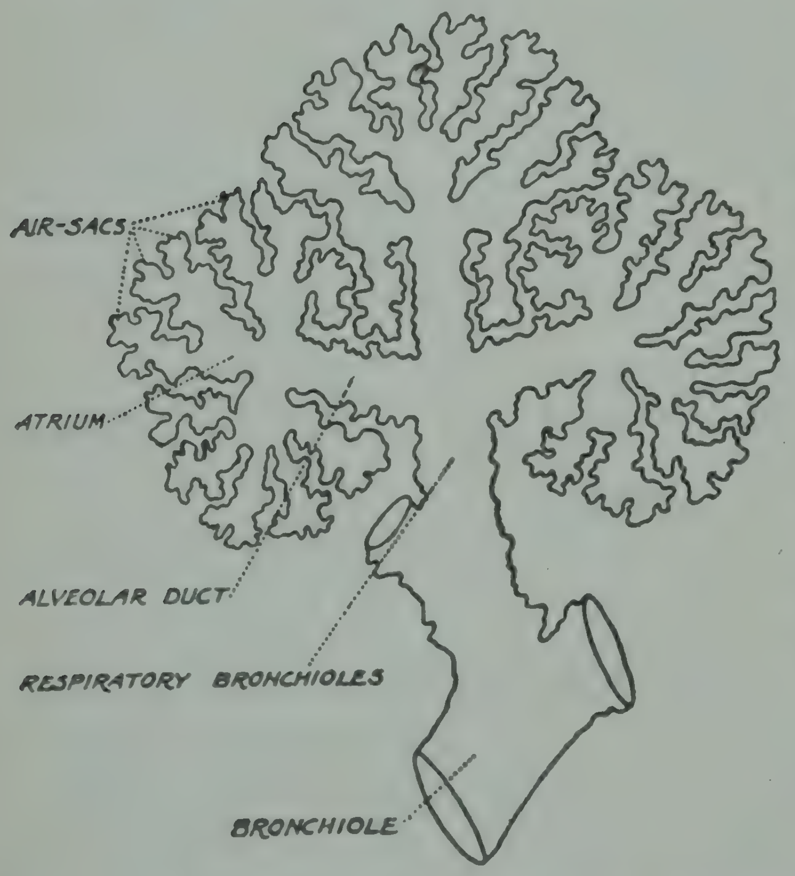


Fig. 477. Diagram of the Air-chambers derived from a Bronchiole. (After Miller.)

The trachea divides below into the two main *bronchi*, and these subdivide into primary and secondary bronchi, etc., each branch becoming smaller, although the total area of their cross-section is together greater than that of the parent tube. The terminal ramifications of the conducting airway, called *bronchioles*, open into *respiratory bronchioles*, and these branch again into *alveolar ducts*.

Confusion has arisen in the use of the terms " atria " and " infundibula," which, in any case, have no particular physiological significance ; it is proposed therefore, to omit these, and to regard the alveolar ducts as forming a branching network of tubes, opening into *alveolar sacs* or air sacs, in the walls of which are situated minute sacculations, the *alveoli*. A few alveoli also open from the walls of the ducts and respiratory bronchioles (Fig. 477).

The larger tubes are kept patent by incomplete rings of cartilage which ar reduced to irregular plates in the smaller tubes and disappear in the bronchioles The walls contain fibrous and elastic tissue throughout, while in the conductin tubes there is a muscular layer which increases down to the terminal bronchioles, where it is thickest in proportion to the other layers. Thence it continues, with the elastic tissue, to invest the respiratory portions of the lung, though much reduced in quantity, being broken up by the alveoli which are evaginated through its meshes.

The lining epithelium is ciliated and columnar in the larger tubes, changing to cubical in the respiratory bronchiole, beyond which it becomes flattened to the " pavement " type which lines the alveoli.

The walls of the airway are supplied with *arterial blood* by the *bronchial arteries*, and the respiratory parts of the lung receive venous blood from the pulmonary arteries.

The alveoli are the special respiratory parts of the lung. Their thin walls are composed of connective tissue containing a large number of elastic fibres, and are covered internally by a single layer of extremely thin, large, flattened epithelial cells. The alveoli are closely packed together, so that, in a section of the lung, an alveolus is seen to be in contact with others on all sides. The alveoli are not blind sacs, but communicate freely with adjoining alveoli by means of " vents " or stomata in their walls.* These probably serve to equalise air pressures in different alveoli, and to enable the alveoli still to be filled if the duct is occluded.

Immediately outside the pavement epithelium ramify blood-capillaries derived from the pulmonary artery. These form a close network, and the venous blood in them is in proximity to air on two sides, being separated from the air in the alveoli only by the thin endothelial cells of the capillary wall and the flattened cells lining the alveoli. It is customary to speak of this barrier between air and blood as the *alveolar membrane*, and it is through this membrane that the gaseous interchange occurs.

For Reference

MILLER (1924). Normal and Pathological Histology of the Lungs. THE HARVEY LECTURES, XX, p. 42. (Philadelphia.)

THE NASAL PASSAGES. The complex nasal passages are covered with a very vascular mucous membrane, by which the incoming air is warmed, and since the mucosa is endowed with mucous glands, the air is also moistened and freed from dust.

It is common knowledge that the nasal passages vary considerably from time to time as regards the ease with which they allow air to enter. Apart from a condition of œdema of the mucosa due to infection, the patency of the nasal passages is subject to change by reason of the variable engorgement of its submucous blood vessels.† The degree of patency may be estimated by blowing a constant stream of air into a nostril, and measuring the pressure of the entering stream ; if the passage is constricted, the pressure is raised, and *vice versâ*.‡ Spraying of the nasal mucosa with adrenaline constricts the vessels and lowers the resistance to air entry ; histamine has the reverse action. Exposure of the body to heat, or to infra-red rays causes closure, while short waves, or skin-cooling, opens the passages. The explanation of the phenomenon is disputed, but it appears to be due to a reflex from the skin. The following table shows the effect of various conditions on the cross-section of the nasal passage, as measured by the rhinomanometric method mentioned above, on a large number of subjects.

AVERAGE CROSS-SECTION IN MM^2 OF NASAL PASSAGES, AT PRESSURE OF 40 CM. WATER (VAN DISHOECK)

Initial cross-section	Electric fire for 10 minutes	Histamine 1 : 1000	Adrenaline 1 : 1000
20·8	16·8	10·2	25·1

The effects of stuffy rooms with central heating bears out these observations.

---

* MACKLIN. *J. Anat.*, 1935, **69,** 188.

† L. HILL. *J. Heat. and Vent. Eng.*, 1934 (March). DUFTON and BEDFORD. *J. Hyg.* 1933, **33,** 476.

‡ VAN DISHOECK. *J. Hyg.*, 1935, **35,** 185.

Respiratory Movements. A constant renewal of the air in the lungs is secured by movements of the thorax; this constitutes normal breathing. With inspiration the cavity of the thorax is enlarged, and the lungs swell up to fill the increased space. The capacity of the air passages of the lungs being thus increased, air is sucked in through the trachea. Inspiration is immediately followed by expiration, which causes diminution of the capacity of the thorax and expulsion of air. Expiration normally lasts 1·3 to 1·4 times the length of inspiration. Expiration is followed, normally without a pause, by inspiration. The number of respirations in the resting adult is commonly about 17 or 18 a minute. This is, however, much influenced by various conditions of the body and also by age. Thus a newborn child breathes about 14 times a minute, a child of five about 26 times, a man of twenty-five about 16, and of fifty about 18. The frequency is increased by any muscular effort, even that of standing. These movements are much affected by psychical activity; they are to a certain extent under the control of the will, although they can occur in an animal deprived of its cerebrum, and are normally carried out without any special act of volition. We can breathe fast or slowly at pleasure, and can even cease breathing for a time. It is impossible, however, to prolong this respiratory standstill for much more than a minute; the need of breathing becomes imperative, and against our will we are forced to breathe.

With every inspiration, the cavity of the thorax is enlarged in all dimensions, from above downwards by the contraction of the diaphragm, and in its transverse diameters by the movements of the ribs. Quiet breathing is almost entirely diaphragmatic.*

The Diaphragm consists of a central tendon which forms an arched double cupola, to the circumference of which are attached muscle fibres. The diaphragmatic muscles present two main divisions, namely, (1) the spinal or crural part, the fibres of which arise from the upper three or four lumbar vertebræ and from the arcuate ligaments, and are inserted into the posterior margin of the central tendon; and (2) the sterno-costal part, which arises by a series of digitations from the cartilages and adjoining bony parts of the lower six ribs and from the back of the ensiform process. These latter fibres pass backwards as they ascend. In the cavity of the larger dome on the right side lies the liver, while the smaller dome on the left side is occupied by the spleen and stomach (Fig. 478). These viscera, in the normal condition, are pressed against the under surface of the diaphragm by the elasticity of the abdominal walls. The central part of the diaphragm is thus pressed up into the chest, partly by the intra-abdominal pressure, and partly by the elastic traction of the extended lungs. The upper surface of the central tendon is united to the pericardium. This part, during expiration, is the highest part of the middle portion of the diaphragm. In expiration, the lateral muscular zone of the diaphragm lies in contact with the lower part of the thoracic wall (Fig. 478). During inspiration, the muscle fibres contract and draw the central tendon downwards, so that the lower surface of the lungs descends. The enlargement of the lungs at the lower part of the thorax is aided by the abduction of the floating ribs, produced by the contraction of the *quadratus lumborum* and *deep costal* muscles. In this contraction the diaphragm presses on the contents of the abdomen, so that the abdomen swells out with each inspiratory movement. The middle of the central tendon, where the heart lies, moves less than the two domes, and the part where the vena cava passes through the tendon is practically stationary during normal respiration. In deep inspiration, however, both this part as well as the rest of the pericardial attachment is forcibly depressed towards the abdomen. In quiet breathing, when observed by X-rays, the mean descent of the right dome in inspiration has been found to be about 12·5 mm., and of the left dome 12 mm. We may say, roughly, that the average descent of the diaphragm during normal respiration is about half an inch. The viscera and the intra-

---

* Hoovre. *Arch. Int. Med.*, 1922, **30**, 1.

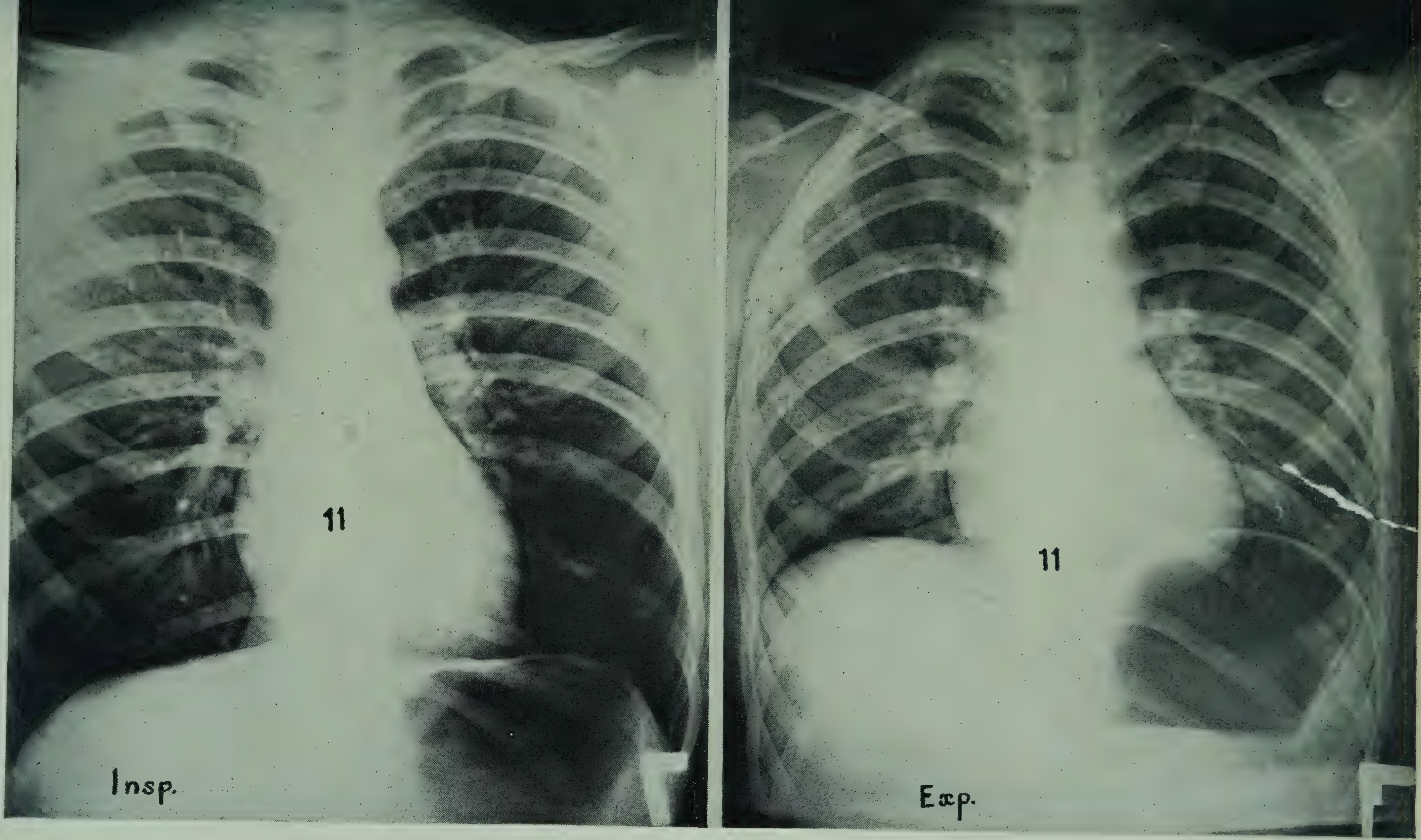


FIG. 478. Postero-anterior radiographs (negatives) of the thorax of a youth aged 18, taken at 6 feet distance. The level of the 11th thoracic vertebra is indicated as a landmark. In full inspiration note the darkness (*i.e.* transparency) of the lower half of the lung-fields, the width of the intercostal spaces and the outline of the heart, diaphragm, root of lung, and the gas-filled fundus of the stomach. In full expiration note the lessened opacity of lower half of the lung-fields, the obliquity of the ribs, the altered position of the heart, root of lungs and diaphragm, and the expansion of the gas in the stomach fundus. (Radiograph specially made by DR. JAMES F. BRAILSFORD.)

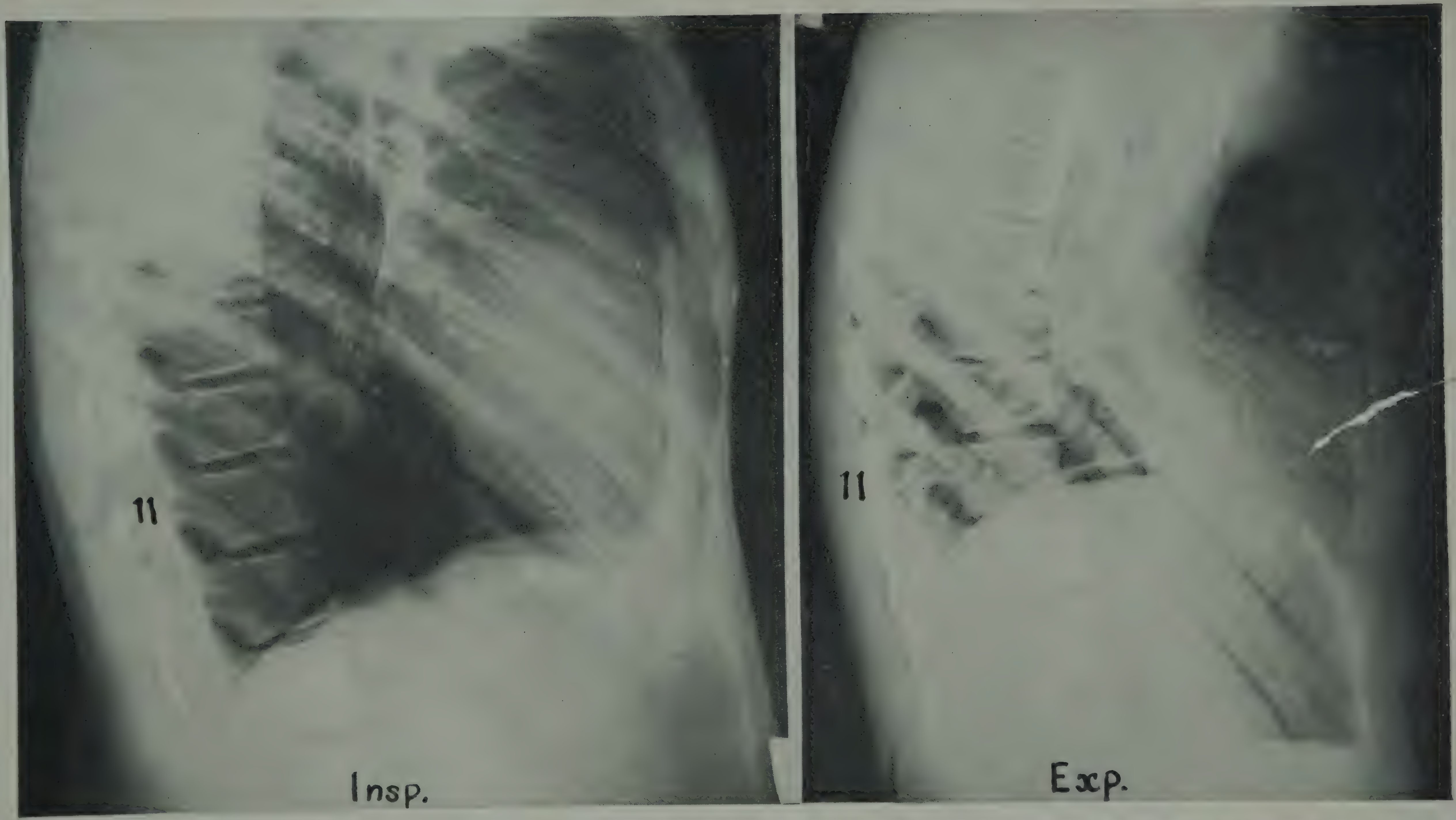


FIG. 479. Lateral radiograph of the chest of the same subject as the preceding figure. Note particularly the reduced obliquity of the ribs and the darkness of the lower lung-fields in full inspiration. (Radiograph specially made by DR. JAMES F. BRAILSFORD.)

abdominal pressure play an important part in determining the movement of the diaphragm, and especially in preserving the abduction of the lower ribs and so furnishing a fixed point for the muscular fibres of the diaphragm. If the contents of the abdomen are removed from a living animal, the ribs are drawn inwards every time the diaphragm contracts. In children with weak chest walls, and with respiratory obstruction, we may often see a depression round the lower part of the chest, which corresponds to the line at which the diaphragm leaves the chest wall, so that the distending force of the abdominal pressure on the bony walls of the thorax abruptly gives place to the pull of the distended lung. The contraction of the diaphragm may be regarded as a rhythmic fluctuation of tonus ; in quiet breathing, at all events, only some of the fibres, and not always the same ones, "contract," which explains the absence of fatigue.

Rib Movements. The enlargement in the other diameters is effected by an elevation of the ribs (Figs. 478 and 479). Each pair of corresponding ribs forms a ring directed obliquely from behind downwards and forwards. With each inspiratory movement the ribs are raised and the horizontal distance between sternum and spinal column is therefore increased. Moreover, the ribs from the first to the seventh increase in length from above downwards, so that when they are raised, the sixth rib, for instance, occupies the situation previously taken by the fifth, and the transverse diameters of the thorax at this level become greater. With each inspiration, there is also a rotation of the ribs. In the expiratory condition their outer surfaces are directed not only outwards but also downwards. As they are raised by the inspiratory movements, they rotate on an axis directed through the fore and hind ends of the rib, so that their outer surfaces are turned directly outwards. In this way a certain enlargement of the thoracic cavity is produced. As the thorax is raised, there is always some twisting of the rib cartilages.

The chief muscles raising the ribs in inspiration are the external intercostals ; their action will be explained later. The serrati postici and scaleni also contribute. It is probable that an important part is played by the extension of the spinal column. This movement, which is specially marked at the upper part of the thorax, causes an increase in all three diameters of this cavity.

The muscles mentioned are the only ones normally engaged in carrying out inspiration. When, from any cause, the inspiratory efforts become more forcible, a large number of accessory muscles are brought into play. These are : the scaleni, sterno-mastoid, trapezius, pectoral muscles, rhomboids, and the serratus anticus.

The force exerted by the inspiratory muscles is nearly all spent in overcoming the elastic resistance of the lungs and costal cartilages. A free access of air is provided for by contractions of certain accessory muscles of respiration. Thus there is rhythmical separation and approximation of the vocal cords, synchronous respectively with inspiration and expiration (Fig. 315, p. 467). When inspiration is laboured, the alæ nasi are dilated by the action of the dilator nasi. This movement, which is constant in many animals, becomes very marked in children suffering from certain types of respiratory trouble.

Expiration. In expiration, the processes are reversed, and the cavity of the thorax is diminished in all three dimensions. Normal expiration is chiefly effected passively. When the inspiratory muscles cease to contract, the stretched lungs contract by virtue of the elastic tissue they contain, and the thorax itself sinks by its own weight, and by the elastic reaction of the stretched costal cartilages.

In a position of rest, however, the elasticity of the thorax is opposed to that of the lungs and would, therefore, tend to produce inspiration. This would make inspiration easier at its onset, and would also present an impediment to expiration, towards the end of which there is need for active muscular contractions. Hence, contraction of the expiratory muscles reciprocates with that of the inpsiratory muscles. In fact, Sherrington's results on the co-ordination of muscular movements would lead us to expect inhibition of the tone, *e.g.* of the abdominal muscles, during inspiration, and active augmentation during expiration. Where the tone of the muscle is entirely lost, *e.g.* in visceroptosis, the diaphragm is thrown out of action, breathing being carried out chiefly by an elevation of the upper part of the thorax. Under normal circumstances, the internal intercostal muscles also contract with each expiration. When the nerves to external and internal intercostal muscles are led off to recording oscillographs, showers of nerve impulses are seen to pass down these, to the external intercostals in inspiration and to the internal intercostals in expiration.* When the breathing is deepened the impulses are seen to be more frequent.

In forced expiration, other muscles may take part—such as the serratus posticus inferior and the muscles forming the wall of the abdomen, *i.e.* the rectus, obliquus, and transversus abdominis muscles.

Although the action of the intercostal muscles has been a subject of debate and cannot even yet be regarded as definitely settled, physiological experiments serve on the whole to confirm the view first put forward by Hamberger, and based on a consideration of the direction of the fibres. The external intercostals pass from one rib to the next below, downwards and forwards. Hence, if a pair of ribs be isolated from the rest of the chest wall, leaving the vertebral and costal attachments intact, contraction of these muscles will cause both ribs to rise.† This result will be evident from a consideration of Fig. 480, where *ab* is a fibre of the external intercostal muscles passing from the rib *vs* to be attached to the rib *v′s′* at *b*. When *ab* contracts, the tension it exerts on its two attachments can be resolved into two components *ac* acting downwards and *bd* acting upwards. *bd*, however, acts at the end of the long lever *bv′*, whereas *ac* acts at the end of the short lever *av*. Hence the raising effect will overcome the depressing effect, and both ribs will rise.

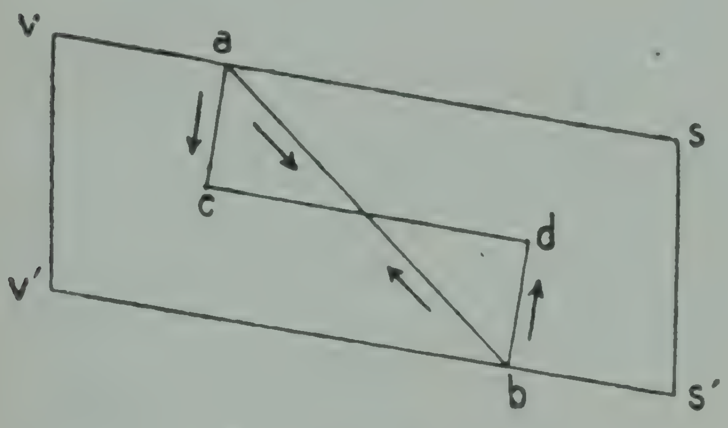


Fig. 480.

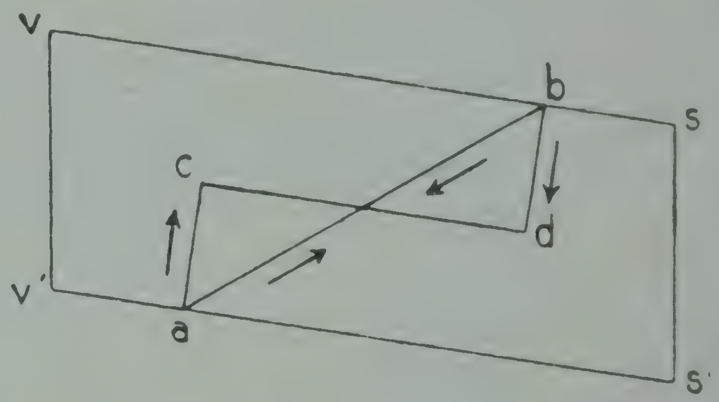


Fig. 481.

The fibres of the internal intercostals run in the opposite direction to the external muscles, and from a consideration of Fig. 481 it is evident that their effect will be to depress any pair of ribs, thus acting as expiratory muscles.

Owing to the fact that the costal cartilages make an angle with the bony ribs, the fibres of prolongation of the internal intercostals, *musculi intercartilaginei,* have the same relation to their attachments that the external intercostals have to the bony ribs. Their action, therefore, must be to raise the cartilages and flatten out the angle between the cartilaginous and bony ribs, so that they must act with the external intercostals as inspiratory muscles.

Lung Movements. As the lungs are inflated their position changes in relation to the thoracic wall. All parts are not equally distensible in the normal position of the lungs. There are three areas which are in contact with the nearly stationary parts of the thoracic wall and cannot therefore be directly expanded. These are (1) the mediastinal surface in contact with the pericardium and structures of the mediastinum ; (2) the dorsal surface in contact with the spinal column and with the spinal segments of the ribs ; (3) the apical surface lying in contact with the deep cervical fascia at the root of the neck. The expansion of these portions of the lung is in a downward direction and is indirectly caused by the movement of the diaphragm. The roots of the lungs move, with inspiration, slightly forwards and downwards. With inspiration there is also an increase in length and width of all the air-tubes, even of the trachea and large bronchi, but especially of the smaller ones. The actual alveoli are probably little, if at all, stretched at inspiration, but merely become flatter and wider. At expiration, all these changes are reversed.

---

* Bronk and Ferguson. *Amer. Journ. Physiol.*, 1935, **110**.

† Sharpey-Schafer and Macdonald. *J. Physiol.*, 1925, **60**, *Proc.* xxv.

Methods for Study of Respiratory Movements. In man, the enlargement of the chest and the accompanying movements of the diaphragm, ribs, heart, root of lungs, etc., can partly be seen on visual, or better on X-ray examination ; records of the expansion of the chest or abdomen can be obtained by the use of stethographs, of which there are many patterns, mostly on the air-transmission principle (Fig. 482).

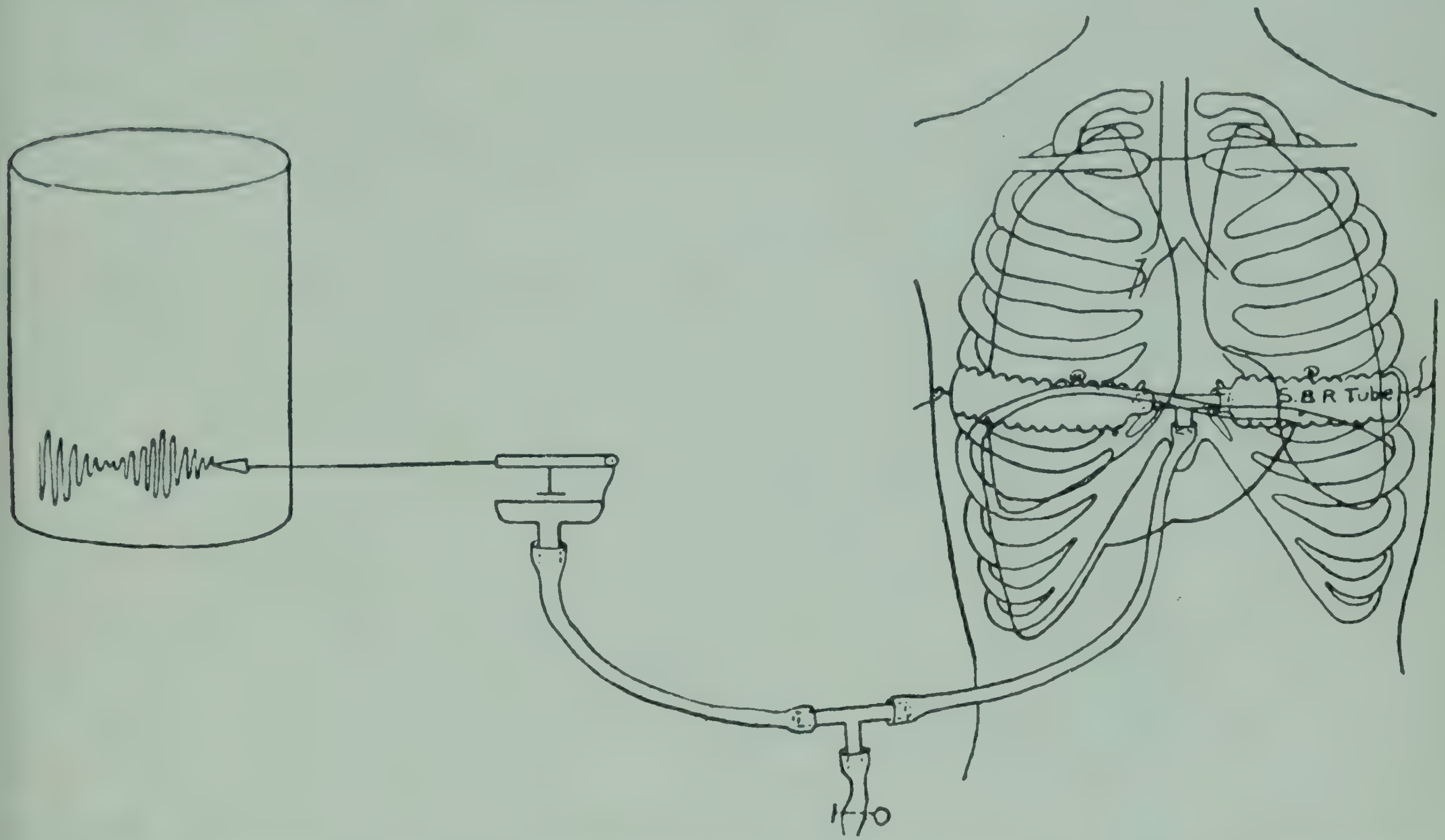


Fig. 482. A convenient form of stethograph. The portion placed round the chest consists of corrugated (" Gorman ") tubing, which opens out like a concertina on inspiration. (Harris's " Experimental Physiology.")

In animals a record may be made of movements of portions of the diaphragm, of the sternum or ribs, or of the larynx. A convenient method is to insert a small rubber balloon filled with air or water beneath the diaphragm and connect it with a water manometer or volume recorder. Yet another method is to record the intrapleural pressure changes,* or the pressure changes in the trachea. The volumes of air breathed in or out may be measured by a gas-meter or spirometer through which the man or animal breathes by means of appropriately arranged valves.† A method giving quantitative records is to place the man, or animal, in an airtight box, from which the head protrudes, the neck passing through an airtight collar : the interior of the box is connected to a volume recorder.‡ Other quantitative methods can be used, but cannot be described here.§

The extent and boundaries of the lungs can be ascertained in the living subject by means of *percussion*. On tapping a finger laid on the chest, a sound is emitted, which varies with the nature of the subjacent tissues. If this is lung tissue filled with air, a clear resonant tone is obtained ; where it is solid tissue, such as the heart, or a lung consolidated with inflammatory products, or the liver, a dull sound is obtained. It is easy to show that the resonant area of the chest increases with each inspiration. The apices of the lungs extend about one inch above the clavicle anteriorly, and reach as high as the seventh spinous process behind. During moderate expiration the lower margin of the lungs extends, in front, from the upper border of the sixth rib at its insertion to the sternum, and runs obliquely downwards to the level of the tenth rib at the back of the chest. During the deepest inspiration the lungs descend, in front, to the seventh intercostal space, and behind, to the eleventh rib, while during deepest possible expiration, the lower margins of the lungs are elevated almost as much as they descend during inspiration. In the front of the chest a triangular space can always be marked out over the heart, where the note obtained on percussion is dull. This space is bounded on the right by the left border of the sternum and extends out as far as the cardiac apex, being bounded above by the fourth costo-sternal articulation and below by the sixth costal cartilage.

---

* Franklin and Gilding. *Journ. Physiol.*, 1932, **75**, *Proc.* i.

† Douglas and Priestley. "Human Physiology." Oxford, 1924. Lamb. "An Introduction to Human Experimental Physiology." London, 1930.

‡ Haldane and Priestley. *Journ. Physiol.*, 1904, **32**, 242.

§ Gaddum. *Journ. Physiol.*, 1941, **99**, 257.

Breath Sounds. If the ear be applied to the chest wall, either directly or through the medium of a stethoscope, each inspiration is found to be accompanied by a fine rustling sound, the "vesicular murmur." It is thought to be caused by the sudden dilatation of the air vesicles during inspiration, or perhaps by the current of air passing from the terminal bronchioles into the wider respiratory bronchioles. It is important to remember that this sound is heard only during inspiration, and over healthy lungs. On listening over the larger air passages, *i.e.* the larynx, trachea, and bronchi, we hear a much louder sound which accompanies both expiration and inspiration, and may be compared to a sharp whispered *hah*. This is known as the "bronchial murmur." It can be heard also at the back of the chest between the scapulæ at the level of the fourth dorsal vertebra, where the trachea bifurcates. In all other parts of the chest the healthy lung prevents the propagation of this sound to the chest wall. If, however, the lung is solid, as occurs in pneumonia, it conducts the sound easily from the large air tubes to the chest wall. Bronchial breathing at any part of the chest, other than that immediately over the air tubes, is therefore a sign of consolidation or cavitation of the lung. Absence of breath sounds at any part of the chest implies, either that air is not entering that part of the lung, or that the lung is separated from the chest wall, *e.g.* by effused fluid, which does not conduct the sound to the surface.

Intrathoracic and Intra-abdominal Pressures. Even at the end of normal expiration the healthy adult lungs are in a stretched condition. This is shown by the fact that if an opening be made into the pleural cavity, air rushes into the opening (*pneumothorax*)* and the lungs collapse, driving a certain amount of air out through the trachea. Since the lungs are always tending to collapse, it is evident that they must exert a pull on the thoracic wall. The pressure in the pleural cavity, between the two layers of the pleura, is called the "intra-pleural pressure." Owing to the pull of the lungs it is below the pressure of the surrounding atmosphere and is sometimes, though inaccurately, called a "negative pressure" in consequence. It is often imagined, though quite wrongly, that the existence of this "negative pressure" is in some way essential to the performance of inspiration, in which act it becomes still more "negative." Actually, of course, the pleura covering the surface of the lungs is in contact, except for a film of pleural fluid, with the pleural layer lining the chest wall, so that the "pleural cavity," in the normal chest, is almost non-existent, and in any case does not contain air. The so-called "negative pressure" is again often, and once more wrongly, stated to be what keeps the two layers of the pleura in contact. So far from being the cause, it is, in fact, an effect, firstly of this contact, secondly of the normally stretched condition of the lungs. The effect of the contact between the two pleural surfaces is that, in order to separate them, the film of liquid by which they are wetted must be broken. To do this requires considerable force, a film of water at body temperature being able to sustain a tension of 3,600 mm. Hg/sq. cm. without rupture.† Everyone knows that what makes a dry rubber sucker adhere is not suction within but pressure of air from without; remove this air pressure and the sucker fails to stick. But a wet sucker would still adhere unless pulled sufficiently to break the fluid film. The lungs, therefore, cling to the chest wall because they fill all the available space of the thorax. If, in so doing, they are stretched, there will be a tendency, proportional to the stretch, for the lungs to collapse, and this will cause the so-called negative pressure. If the atmospheric pressure inside the lung $= P$, and the elastic recoil of the lungs $= p$, then the two layers of pleura will be held together by the force $P - p + S$, the "negative pressure" of course also being $= p$, and exerted on the whole contents of the thorax, and S being the initial force required to break the liquid film. In the expiratory position the intrapleural pressure is 3 to 5 mm. Hg below the atmospheric; in normal inspiration the difference is about 5 to 10 mm. below atmospheric. If the lungs are fully distended by a deep inspiration the elastic forces are brought more into play, and the negative pressure in the pleura may amount to 30 mm. But it must be clearly understood that if the lung tissue had no elastic recoil, so that there was no "negative pressure," the act of inspiration would not be affected in any way. But respiration becomes impossible directly free openings are made into the pleural cavities on both sides of the chest. With each inspiratory movement, air rushes in through these openings, so that the thoracic movements can no longer exert any influence on the volume of the lungs.

---

* Graham. Alterations of Intrapleural Pressure and their Significance. Harvey Lectures, XIX, 1924, 123. Lippincott.

† Burns. *J. Physiol.*, 1940, **98**, 26P.

The negative pressure in the thorax is diminished by any factor decreasing the elasticity of the lung tissue. Thus, in an old man, where the elastic tissue is degenerated and the alveoli are enlarged, giving rise to the condition known as *emphysema*, the lungs may collapse only slightly, or not at all, on opening the chest. The lungs do not collapse on making an opening in the chest of a new-born mammal; but this is owing to the fact that they completely fill the thorax in the expiratory position, and it is only later that, with the growth of the ribs, the thorax gets, so to speak, too large for the lungs, which are therefore stretched to fill it.

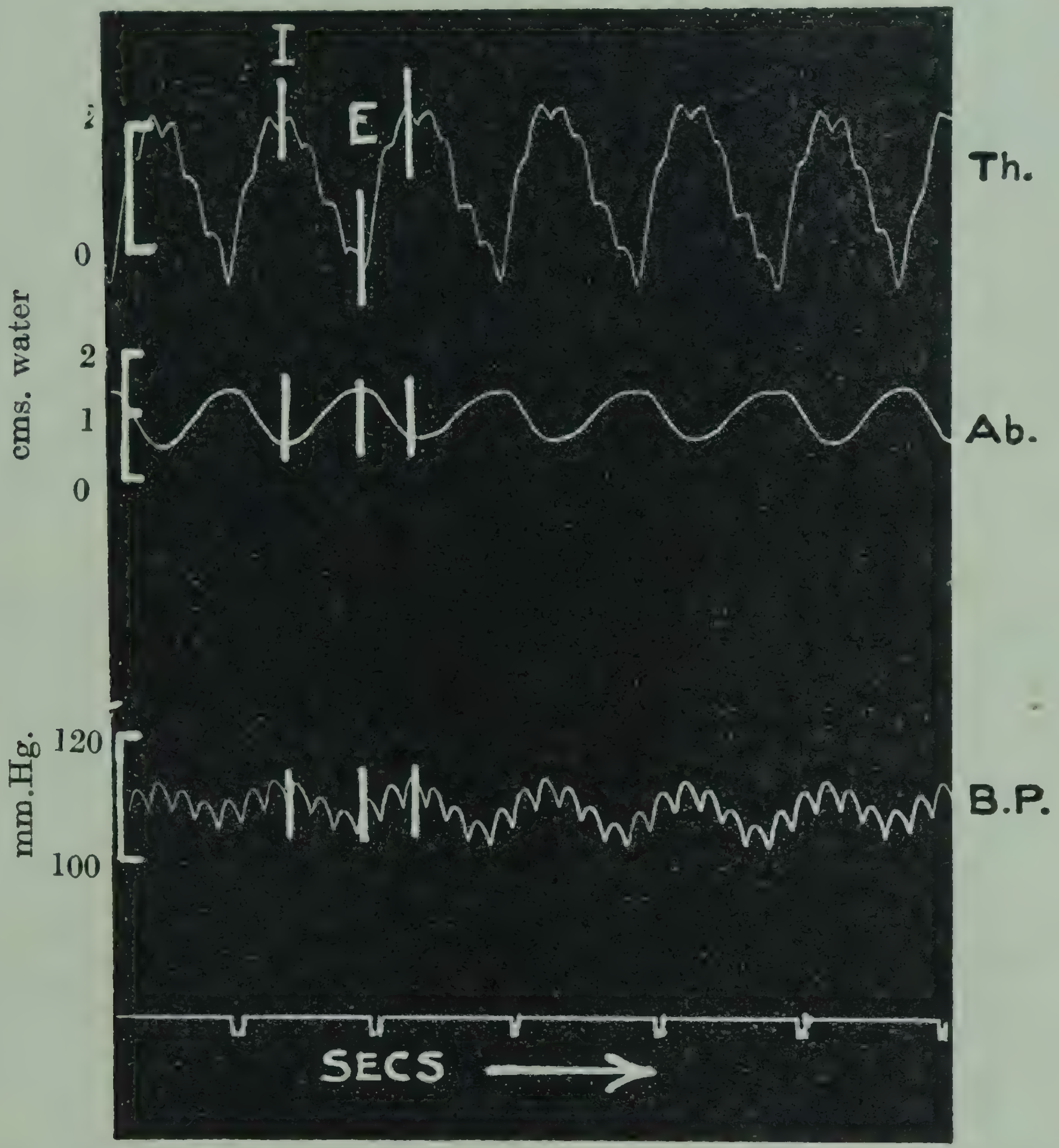


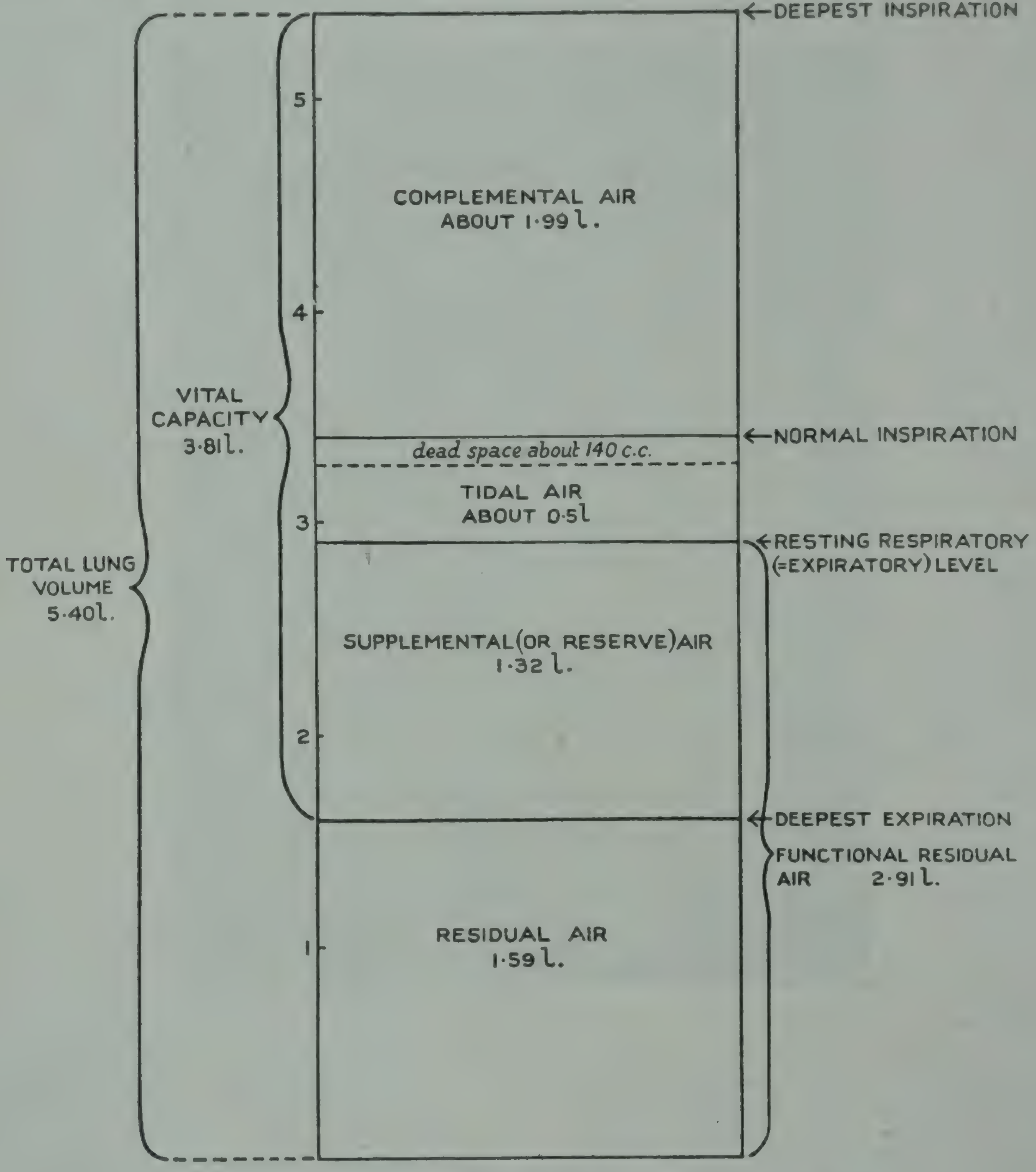
Fig. 483. Fluctuations of intrathoracic (Th) and intra-abdominal (Ab.) pressure (in terms of cm. water) and of arterial blood pressure (in mm. Hg) of an anæsthetised cat breathing quietly. I = inspiration begins. E = expiration begins. (H. P. Gilding.)

The pressure in the abdominal cavity is of course raised when the diaphragm descends during inspiration, so that the pressure fluctuations in the thoracic and abdominal cavities are reciprocal in direction (Fig. 483). Whether the actual pressure at any point in the abdominal cavity is above or below that of the outside air will obviously depend greatly upon the position considered, the posture of the body, the tension on the abdominal walls, etc., but in any cases with normal breathing, the pressure in the abdomen will rise, as that in the thorax falls, during inspiration, and *vice versâ*. This is a fact of the greatest importance in several directions.

If a manometer be connected with one of the nostrils, so as to register the pressure in the air cavities, it is found that there is a negative pressure of

— 1 mm. Hg with inspiration, and a positive pressure of 2 or 3 mm. with expiration. With forced inspiration, the pressure may amount to — 57 mm. Hg, and with forced expiration there may be a pressure of + 87 mm., the latter mainly due to action of the abdominal muscles.

Pulmonary Ventilation. When breathing quietly, a man takes in and gives out at each breath about 500 c.c. of air, measured dry and at 0° C. If measured moist and at the temperature of the body, viz. 37° C., the volume would be about 600 c.c. This amount is known as the *tidal air*. By means of a forcible inspiratory effort, it is possible to take in about 2,000 c.c. more (*complemental air*). At the end of a normal expiration a forcible contraction of the expiratory muscles will drive out about 1,300 c.c. more (*reserve or supplemental air*). These three amounts together constitute the "*vital capacity*" of an individual. This total may be determined by means of the spirometer, a small gasometer with a gauge by which the amount of air in it can be at once read off. The person to be tested fills his lungs as full as possible, and then expires to the utmost into the spirometer. Individual variations in vital capacity are great (2·5—5·5 l.), and depend to a considerable extent on physical build. Of the total lung volume, the vital capacity of British subjects averages about 71 per cent., supplemental 25 per cent., functional residual air 54 per cent., residual air 29 per cent.* The *functional residual air* is supplemental plus residual air (Fig. 484). The pulmonary ventilation and vital



DEEPEST INSPIRATION
5
COMPLEMENTAL AIR
ABOUT 1·99 l.
4
VITAL CAPACITY 3·81 l.
NORMAL INSPIRATION
dead space about 140 c.c.
TIDAL AIR
ABOUT 0·5 l
3
RESTING RESPIRATORY (=EXPIRATORY) LEVEL
TOTAL LUNG VOLUME 5·40 l.
SUPPLEMENTAL (OR RESERVE) AIR
1·32 l.
2
DEEPEST EXPIRATION
FUNCTIONAL RESIDUAL AIR 2·91 l.
1
RESIDUAL AIR
1·59 l.

Fig. 484. Lung volumes. (Data from Aslett *et al.*)

---

* Aslett, D'Arcy Hart and McMichael. *Proc. R. S.*, 1939, **126B**, 502.

capacity are greatest when standing, and least when lying in the prone position, the difference in V.C. being about 300 to 500 c.c. The loss of vital capacity when lying down is due to upward displacement of the diaphragm by the abdominal viscera, and also, according to Hamilton and Morgan,* to the fact that blood accumulates in the lungs. In cardiac disorders with weakness of the left ventricle, patients often are unable to lie down because of the pulmonary congestion and breathlessness which ensue (orthopnœa).

Increased depth of breathing is called *hyperpnœa.* When for any reason the pulmonary ventilation approximates to the volume of the vital capacity, the breathing becomes laboured, and is called *dyspnœa.* Voluntary hyperpnœa is called *forced breathing.*

Under no circumstances can expiration quite empty the lungs of air. At the end of the most forcible expiration the alveoli and all the air tubes still contain air. The air left in the lungs after the most vigorous expiration is known as the *residual air.* As a result of actual determinations, we may assume the residual air in the lungs as something between 1,000 and 1,700 c.c., generally about 1,600 c.c. It is not reduced in the recumbent posture.

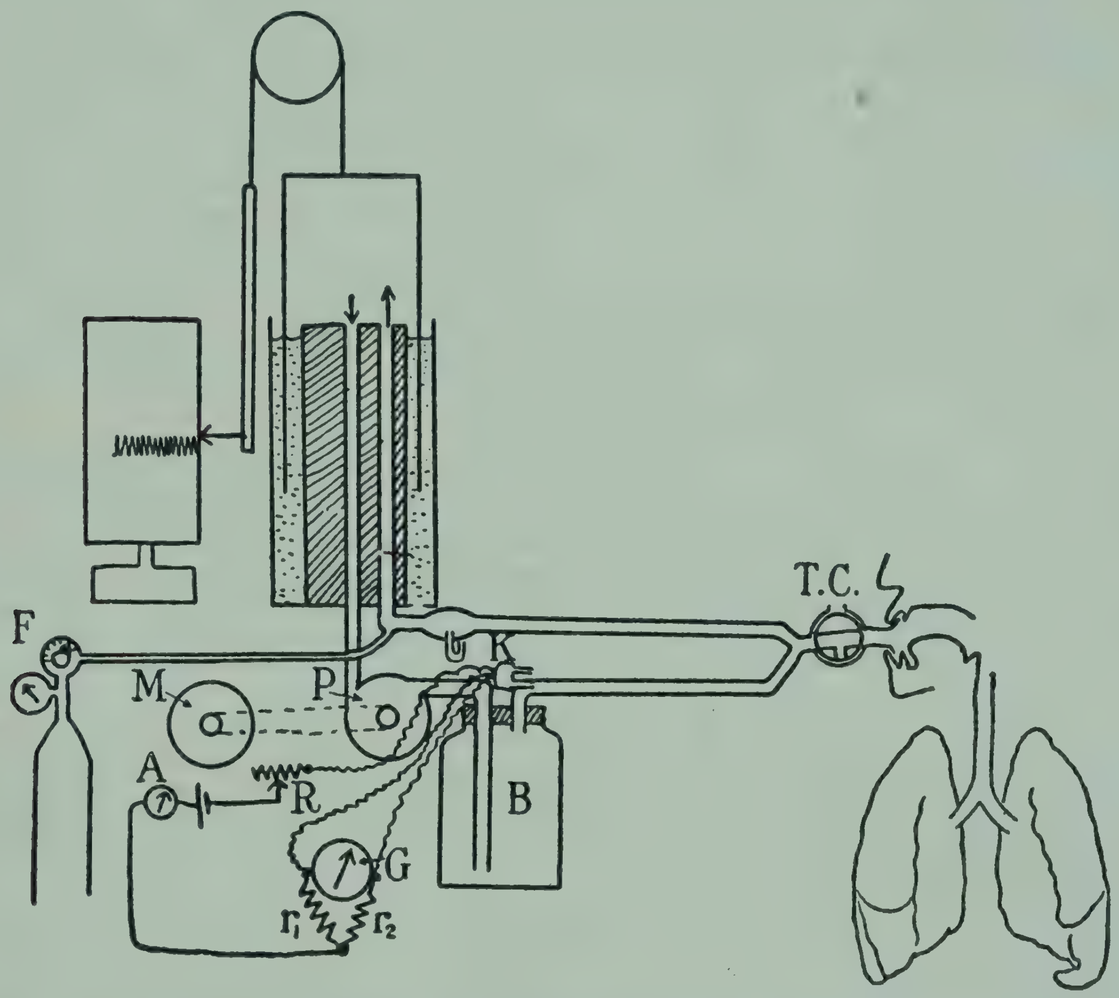


Fig. 485. Spirometer arranged to determine lung volume. P., the circulating pump ; B, bottle with absorbent for CO_2 ; volume kept constant by running in O_2 from F. at same rate as usage ; K, a katharometer connected to ammeter A and galvanometer G, for ensuring constant composition of mixture. (Herrald and McMichael, *Proc. R. S.*, 1939, **126B**, 491.)

The various lung volumes can be accurately determined by the use of a recording spirometer as follows : †

(1) The subject is connected to the spirometer arranged as in Fig. 485 and a record of the rate of oxygen consumption is taken in the usual way. During this time records are also made of *vital capacity, complemental air* and *supplemental air.* The subject is now disconnected.

(2) It now remains to determine *functional residual air.* The principle of the method is to let the patient breathe from an atmosphere of *known volume* and of *known composition,* but differing from that of the air by being greatly enriched with oxygen. The composition of the mixture is kept constant by adding oxygen at the same rate as the subject uses it. The CO_2 is absorbed. In seven minutes the N_2 is evenly distributed, the patient is disconnected and the mixture analysed for O_2 and N_2.

---

* Hamilton and Morgan. *Amer. J. Physiol.*, 1932, **99**, 526.

† Herrald and McMichael. *Proc. R. S.*, 1939, **126B**, 491.

A calculation shows the functional residual air. Thus, suppose the patient's lungs at the start contained x l. of air which = 0·791 x l. of N_2, and suppose the spirometer contains to begin with 2 l. of air (= 2 × 0·791 l. N) plus 3 l. of O_2. The volume is the same at the end because oxygen is added as used, and all that has happened is that the lung gases and spirometer gases have been mixed together. Suppose there is 50 per cent. of N_2 in the final mixture; the functional residual air is x l.

Then we have

$$\begin{aligned} 2 \times 0{\cdot}791 + 0{\cdot}791\,x &= 0{\cdot}50\,(x + 5) \\ 0{\cdot}291\,x &= 0{\cdot}918 \\ \therefore \quad x &= 3{\cdot}16 \text{ l.} \end{aligned}$$

(3) If the supplemental air was found to be, say, 1·56 l., the *residual air* would be 3·16 − 1·56 l. = 1·6 l.

Of the 500 c.c. of tidal air taken in at each inspiration, only a certain part reaches the alveoli, part being required to fill the air tubes, trachea, bronchi, and bronchioles which lead to the air cells. The volume of these air tubes is called the "*dead space*," and is normally taken as about 140 c.c., so that, of the 500 c.c., about 360 c.c. actually reach the alveoli. For the same reason, the expired air represents the air from the alveoli (360 c.c.), diluted with 140 c.c. of dead space air which has remained in the air tubes and undergone very little change, other than the elevation of temperature and saturation with aqueous vapour. There is no doubt, however, that it is incorrect to regard the volume of the dead space as constant, since not only do the branches of the bronchial tree show at inspiration an increase, and at expiration a decrease, in all dimensions, but the volume of the physiological dead space no doubt varies with alterations in the state of the bronchial musculature and various other factors. It is, for instance, increased in exercise.*

The velocity with which the air enters and leaves the respiratory tract may be investigated by photographing the displacement of a fine quartz fibre placed in the path of the moving air. It shows that the maximum speed of movement of the air is at the commencement of inspiration and expiration, and the minimum speed just before the end of each phase.

For Reference

ANDRUS (1936). The Mechanics of Respiration. *Amer. Rev. Tuberc.*, **33**, 139.
HEUSNER (1946). Yawning and Associated Phenomena. *Physiol. Rev.*, **26**, 156.
KEITH (1909). The Mechanism of Respiration in Man. (L. HILL's "Further Advances in Physiology.") London. (Arnold.)

Artificial Respiration. Under various conditions the respiratory movements may cease, and it is then necessary, in order to save life, to cause air to enter the lungs by some artificial means. This may be accomplished in various ways by which air is forced or drawn into the lungs, and forced or drawn out of them, at suitable intervals. The most easily controllable method of doing either is by means of suitable air-pumps, but when these cannot be used, quiet efficient ventilation can be obtained by simple means, as described below.

Inflation Methods. These methods are most commonly applied to anæsthetised animals. In the simplest method, a suitable tracheal cannula is tied into the trachea, and warm moist air is delivered to the tube rhythmically from a pump. The tracheal cannula has a side tube which can be partially closed by a screw clamp; as each blast of air is delivered, some excess escapes by the side tube, and the rest of it inflates the lung; between the blasts the lungs collapse and the air expelled escapes by the side tube. When ventilation is being carried out with gases other than atmospheric air, this method is wasteful. In such cases it is better to use a more elaborate pump, such as the one designed by Starling ("Ideal" pump). In this the side tube is closed by a mechanical device while the pump is delivering, so that the whole of the air delivered is kept in the closed system of lungs plus tubes; at the end of the stroke the side tube is opened and the lungs collapse. In a third type of pump, the trachea is connected to a double pump which alternately forces air into the trachea from one barrel and draws it out again into the other one. The last two methods both allow of respiration being carried out with any desired gas mixture. With suitable modifications in the way of masks instead of tracheal cannulæ, similar devices have been applied for use on the human subject. It was found by Yandell Henderson, Drinker and others that the administration, by some such method, of oxygen containing 5 per cent. carbon dioxide is most valuable in the resuscitation of persons apparently drowned or suffering from coal gas poisoning, and various apparatus for the purpose have been introduced.

---

* DOUGLAS and HALDANE. *J. Physiol.*, 1912, **45**, 235.

Continuous Insufflation. If a flexible tube of small diameter be passed down the trachea as far as its bifurcation, it is possible to maintain life without any movement of the chest at all, by injecting a constant stream of oxygen, with or without 5 per cent. CO_2, by way of the tube, so that the lungs are slightly distended. This has enabled operations to be carried out on patients with opened thorax.

*Chamber Method.** In this method the subject is placed with his body in a chamber and the head outside, the neck passing through an airtight collar in the wall. The chamber is airtight, and when air is sucked out of it by a rhythmically operated pump, the subject's chest is expanded and air is drawn into his lungs ; the reverse happens when air enters the chamber again.

The Bragg-Paul Pulsator † is a simple mechanical device which can be operated by a motor, or by water-pressure for indefinitely long periods.‡ It consists of a rubber air-bag wrapped round the chest and rhythmically inflated and deflated.

Emergency Methods. In animals rhythmic compression of the chest is often sufficient. In man, several adequate methods are available.§ That described by Schäfer,|| is generally adopted for the resuscitation of asphyxiated persons. The subject is placed in the prone position, with the head slightly to one side ; the mouth is cleared and the tongue pulled well forward. The operator, kneeling either at the side, or astride, of the subject, places the palms of his hands flat on the subject's trunk, just over or below the lowest ribs and then by leaning forward presses *gently* forwards and upwards so as to exert pressure there, for, say, two seconds (Fig. 486). He then leans back a little to release the pressure for two seconds and then again repeats the forward movement, and so on. In this way air is forced out of the chest when pressure is applied, because the diaphragm is pushed up ; when the pressure is released, the elastic recoil draws air into the chest. It is easy by this method, *with which every person of adult age should be familiar*, to maintain normal ventilation in an unconscious subject for long periods without undue fatigue on the part of the operator.

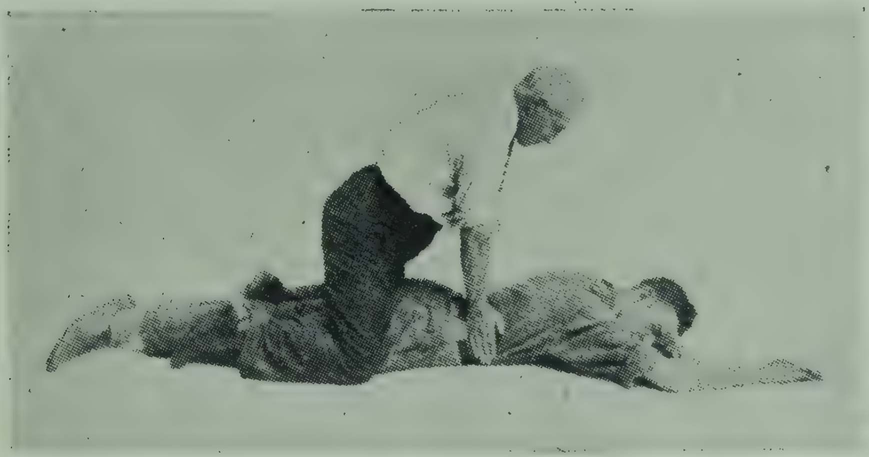


FIG. 486. Schäfer's method of artificial respiration. (DRINKER.)

With patients under operation, and lying in the supine position, this method cannot readily be used ; but quite good results can be obtained by similarly applying intermittent pressure on the epigastrium, by which means also the diaphragm is made to ascend. *Pressure on the upper ribs is almost useless.* In both of these methods it is advantageous, where possible, to let the subject inhale, from a suitable mask, a mixture of oxygen and 5 per cent. carbon dioxide.

Another simple method which can be carried out for long periods is the rocking method of Eve.¶ The patient is placed in the prone position on a

---

* DRINKER and MCKHANN. *J. Amer. Med. Assoc.*, 1929, **92**, 1658.

† PAUL. *Proc. Roy. Soc. Med.*, 1935, **28**, 38.

‡ TOOKEY KERRIDGE. *Lancet*, 1934, *i.*, 786.

§ MACINTOSH and MUSHIN. *B.M.J.*, 1946, *i.*, 908.

|| SCHÄFER. *Trans. Med. Clin. Soc. Lond.*, 1904, **87**, 609 ; *J. Amer. Med. Assoc.*, 1908, **51**, 801.

¶ EVE. *Lancet*, 1932, *ii.*, 995 ; *B.M.J.*, 1943, *i.*, 535.

stretcher, which is rocked on a pivot, like a see-saw, 45 degrees up and down 12 to 15 times a minute. When the head is downwards the weight of the abdominal viscera pushes the diaphragm up, and in the feet-down position the diaphragm descends. The tidal air by this method is slightly greater than in Schäfer's method,* viz. 450 to 600 c.c./min.: it also favours venous return to the heart, and hence the cardiac output is improved.

THE BRONCHIAL MUSCULATURE

All the air tubes have in their walls a middle layer of plain muscle. When the lung is normally distended, this muscle is neither circular nor longitudinal, but forms an oblique network, which by its constriction narrows the lumen and decreases the length of the tube at the same time. The muscle only appears circular and compact when microscopically examined in the collapsed state (Macklin). The guinea-pig is perhaps the only animal in which the bronchial muscle is definitely circular and also compact,† and in this animal the thickest muscle is in the secondary *bronchi* (just after they arise from the primary); this is the site of occlusion which causes death from asphyxia in anaphylaxis in the guinea-pig. Immediately internal to the muscle coat lies a prominent layer of longitudinally coursing branched elastic fibres, and from this elastic lamina fibres extend amongst the muscle fibres and form, as in the arteries, gut, spleen, etc., an intimate "tissue partnership." In expiration, contraction of this "myo-elastic" system plays an important part in decreasing the size of the whole lung, and in addition decreases the capacity of the dead space by shortening and narrowing the airway. Contraction of the bronchial muscle has the following effects: (1) a constriction of the bronchi and bronchioles; (2) a diminution of the air space of the lung, and therefore of the volume of the lung; (3) an increased resistance to the passage of the air into, and out of, the alveoli.

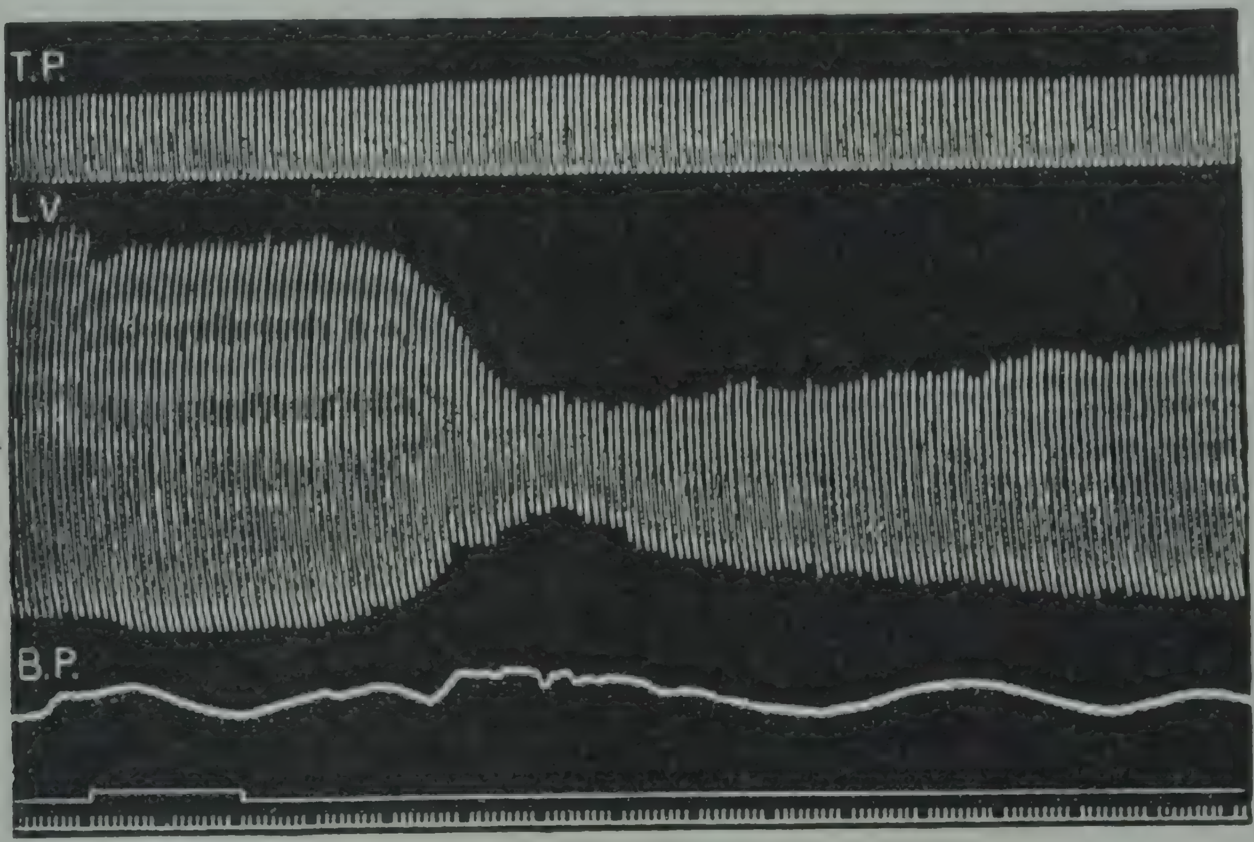


FIG. 487. Tracings of the Volume Changes of the Lung, and of Variations of Tracheal Pressure. (DIXON and BRODIE. *J. Physiol.*, 1903, **29**, 125.)

T.P. tracheal pressure. **L.V.** lung volume. **B.P.** blood pressure (Zero **B.P.** 17 m.m. below time marker). Showing constriction of bronchial musculature as a result of vagus excitation. The increase of **T.P.** is very slight.

---

* KILLICK and EVE. *Lancet*, 1933, *ii.*, 740.

† MILLER. *Amer. Rev. Tub.*, 1921, **5**, 689.

Changes in the condition of contraction of the bronchial muscle may be studied in various ways. In one method, artificial respiration is carried out, a constant volume of air being blown in at each respiration—excess escaping by side-tube, while alterations in pressure due to the varying resistance of the bronchioles may be recorded from a tracheal side tube (Fig. 487, T.P.). Any diminution in the calibre of the bronchioles must increase the resistance to the incoming current of air, and so cause a rise of pressure in the tracheal tube. By a second method, artificial respiration at a constant pressure is made use of, and alterations in lung volume at each pump stroke may be recorded by enclosing the lung in a plethysmograph * (Fig. 487). In a third method, the lungs may be excised and the air tubes perfused from the trachea with a suitable warm saline solution, while the pulmonary circulation is also similarly perfused. The bronchial perfusion fluid flows away through scarifications made on the lung surface, while the resistance is measured by a tracheal side tube and suitable manometer.† The action of drugs or inorganic ions may conveniently be studied by this method. By all these methods it has been shown that stimulation of the peripheral end of the vagus or presence of acetyl choline causes constriction of the bronchioles (Fig. 487) which is paralysed by atropine, while inhibitor fibres supply the bronchial muscle from the sympathetic nervous system ‡ (Fig. 488). The action of adrenaline also causes dilatation of the bronchi. Histamine causes intense constriction of bronchi through its direct action on the muscle,§ and this, like all reactions of plain (and other) muscle, is diminished by reduction of the calcium content of the fluid in which it is bathed.||

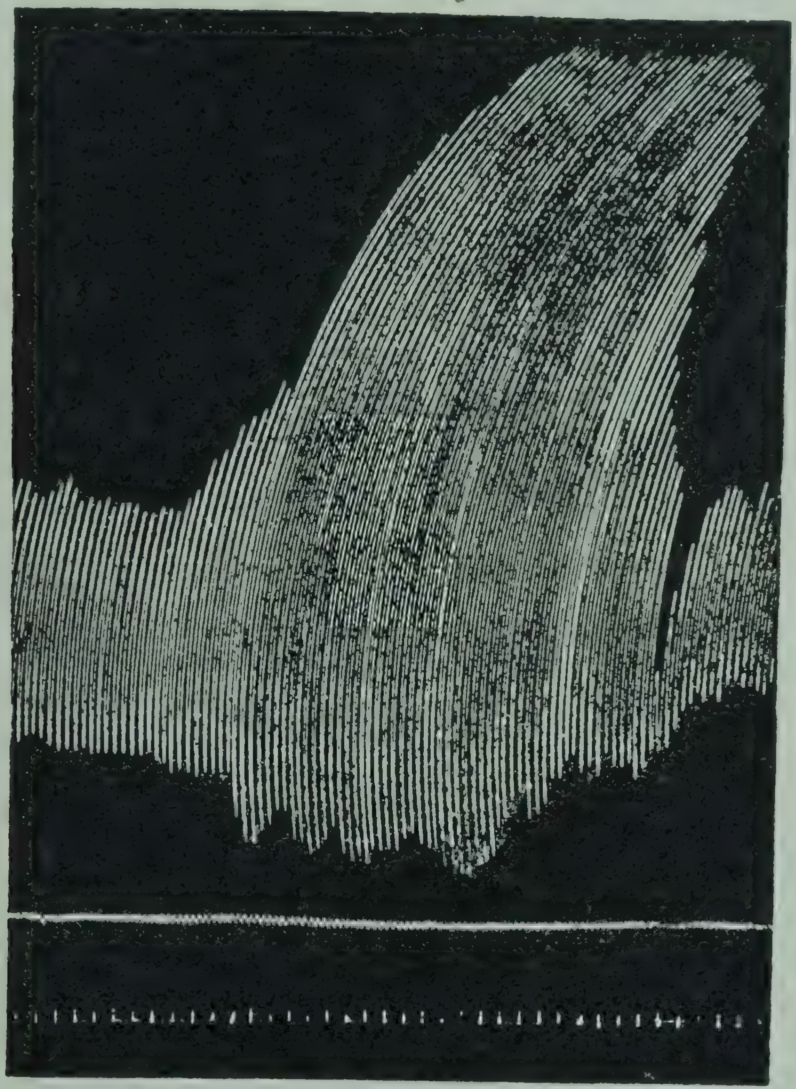


Fig. 488. Effect on Lung Volume of stimulation of post-ganglionic fibres from the ganglion stellatum. Time = 5 secs. (Dixon and Ransom, *J. Physiol.*, 1912, **45**, 413.)

---

* Dixon and Brodie. *J. Physiol.*, 1903, **29**, 97.
† McDowall and Thornton. *J. Physiol.*, 1930, **70**, *Proc.* xliv.
‡ Dixon and Ransom. *J. Physiol.*, 1913, **45**, 413.
§ Thornton. *Quart. J. Exp. Physiol.*, 1932, **21**, 305.
|| Gillespie and Thornton. *J. Pharmacol.*, 1932, **45**, 419.

In man the bronchi have been studied with the aid of the bronchoscope,[*] and by X-rays after the injection of radio-opaque substances ("lipiodol") into the bronchial tree.[†] By both these methods the bronchi have been shown to have a rhythm synchronous with and similar to that of the vocal cords (p. 467), dilating in inspiration and constricting in expiration. The cause of this rhythm is not known. In part it is no doubt passive, and represents a general alternate expansion and contraction of the whole of the lung in consequence of the respiratory movements. But many believe it to be also, in part, active. In so far as this is the case it may be regarded as a part of the normal respiratory movement, caused by the discharge of nervous impulses from the respiratory centre; we may suppose that, at each inspiration, the broncho-constrictor impulses passing down the vagi are inhibited and that broncho-dilator impulses are sent down the sympathetic fibres to the lungs; at expiration the reverse happens.

Peristaltic movements have also been described in the bronchi and are markedly brought into play on the introduction of foreign bodies (Reinberg). (Spontaneous rhythmic waves have been observed in the bronchi of the chick embryo by Lewis.[‡]) These represent a protective mechanism of the lung, which thus raises unwelcome fluid to those parts of the airway (primary and secondary bronchi) where the cough reflex can be originated, by which it is finally expelled.

In asthma, the excessive tonus of the bronchial muscle will cause obstruction to the passage of air in and out of the lung, expiration becoming more difficult than inspiration, probably owing to the rhythm mentioned above. Whether this hypertonus depends on chemical factors or not is at present unknown, but it is clear that nervous elements can be superimposed to add to the difficulties of the already complicated clinical picture.

For Reference

Macklin (1929). The Musculature of the Bronchi and Lungs. *Physiol. Rev.*, **9**, 1.

2. THE CHEMISTRY OF RESPIRATION

Expired Air. The tissues of a man of 70 kg. body weight consume, on an average during working hours, about 400 c.c. oxygen per minute. This oxygen is used for the oxidation of various materials in the body, and the tissues, in consequence, produce 300 to 350 c.c. of carbon dioxide. The oxygen required by the tissues is taken from the blood, and the carbon dioxide formed is, in return, passed out into the blood. Hence, the venous blood contains less oxygen but more CO_2 than arterial blood. In its passage through the lungs, the blood is again arterialized, oxygen passing into it and CO_2 out, owing to an exchange with the air in the alveoli. This air is continually renewed by breathing, and hence the expired air contains less oxygen, and more carbon dioxide than that which is inspired. Its composition is not constant, however.

Expired air can be collected for analysis or measurement by the use of the Douglas bag[§] (Fig. 489). The subject breathes through the valves, which are so arranged that air from the outer atmosphere enters at inspiration, and is drawn into the lungs, while at expiration the expired air passes through the other valve, and into the bag, which is of about 100 l. capacity, and made of an impermeable fabric. Samples of the air can be drawn off from the bag for analysis, and the volume of expired air collected in a given time measured by expressing the contents of the bag through a gas meter.

Analysis of the air is made by the use of Haldane's gas analysis apparatus, of which descriptions will be found in practical handbooks.[||] The principle is that a measured volume of the gas is taken, the CO_2 absorbed by KOH solution, and the shrinkage of volume measured; then the oxygen is absorbed by alkaline pyrogallate, and the volume loss again measured: the residual volume is nitrogen. The whole analysis is conducted at constant temperature and pressure, and with the gas saturated with water vapour. Since only the relative volumes of the components are required, the results are the same as if the gas were dry, and at N.T.P. ; addition of water vapour, or rise of temperature, or fall of pressure, causes expansion of all components equally.

---

[*] Jackson. Bronchoscopy and Œsophagoscopy. 2nd edition, 1927. (Saunders.)

[†] Reinberg. *Brit. J. Radiol.*, 1925, **30**, 451; Bullowa and Gottleib. *Am. J. Med. Sci.*, 1920, **160**, 98.

[‡] Lewis. *Am. J. Physiol.*, 1924, **68**, 385.

[§] Douglas. *J. Physiol.*, 1911, **42**, *Proc.* xvii.

[||] *E.g.* Douglas and Priestley (1948). Human Physiology. Oxford: Clarendon Press.

The expired air of normal resting persons at sea-level contains about 3·0 to 4·5 per cent. of carbon dioxide, and 16·0 to 17·5 per cent. of oxygen. As a very rough average, it compares as follows with the inspired air :

	Inspired Air	Expired Air
Oxygen	20·95 vols. per cent.	16·4 vols. per cent.
Nitrogen (including argon)	79·01 ,, ,,	79·5 ,, ,,
Carbon dioxide	0·04 ,, ,,	4·1 ,, ,,

The inspired air contains a variable amount of aqueous vapour and has a variable temperature. Expired air, as expelled, has a temperature, in our climate, of about 35·5° C. and is about 90 per cent. saturated with aqueous vapour at that temperature, *i.e.* its aqueous vapour tension, as expired, is about 39 mm. Hg. In the alveoli we may assume the air to have a temperature of 37° C. and a tension of aqueous vapour of 45 mm. Hg. Thus when a man is breathing air at a pressure of 760 mm. Hg., the pressure of the mixture of permanent gases in the alveoli of his lungs will be only 760—45, *i.e.* 715 mm. Hg.*

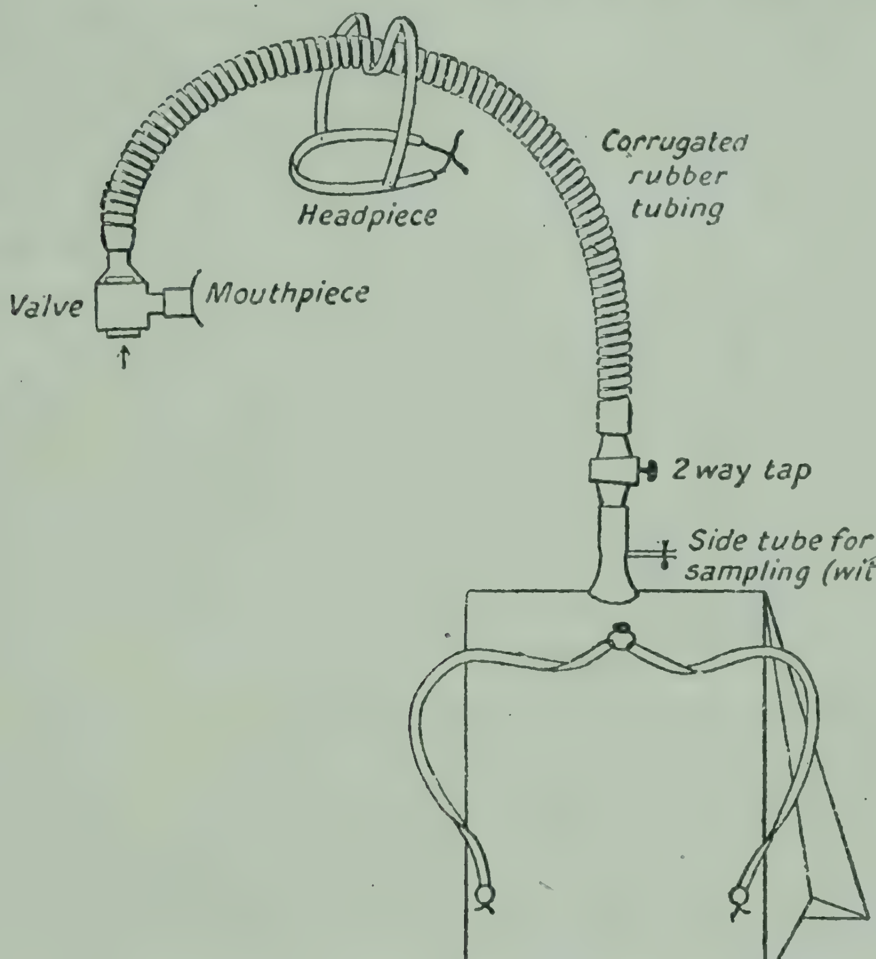


Fig. 489. Douglas' bag, with valves for collection of expired air.

Calculation from these data and from the volume expired shows that, whereas about 250 c.c. of oxygen are used up per minute by a normal resting man, only about 215 c.c. of carbon dioxide are added to the expired air in the same time, while the total *volume* of nitrogen leaving the lungs is the same as that taken in. In the analyses, the nitrogen *percentage* is seen to be higher in the expired air. This is because the volume at N.T.P. of air expired is less than that inspired, owing to the disappearance of a certain amount of oxygen, without the production of a corresponding amount of carbon dioxide, so that the *relative* amount of nitrogen is slightly increased. The reason for this disappearance of oxygen is that when fats and proteins are being consumed, some of it is used for the combustion of hydrogen and for the partial oxidation of nitrogen, as well as in burning carbon. The ratio of the volumes $\frac{\mathrm{CO_2}}{\mathrm{O_2}}$ which are exchanged in a given time, and which is called the *respiratory quotient*, is normally about 0·85, but varies according to the nature of the materials subjected to combustion in the body and to other conditions.

---

* It is generally assumed that the tension of aqueous vapour in alveolar air is 47 mm., *i.e.* that it is in equilibrium with *water* at 37° C. Direct experiment by Christie and Loomis (*J. Physiol.*, 1932, **77**, 35) gives a tension of aqueous vapour of only about 45 mm. Hg in quiet breathing, and still less during hyperpnœa.

Alveolar Air. As already explained, only about 360 c.c. of the 500 c.c. of air drawn in at an average breath reach the alveoli, the other 140 c.c. being dead space air. Hence, the alveolar air must contain more carbon dioxide and less oxygen than the mixed expired air, which consists of alveolar air *plus* the unchanged dead space air, which precedes it on expiration.

At the end of an expiration, the dead space has been swept out by, and remains filled with, alveolar air. Based on this fact, a sample of alveolar air may be obtained for analysis in the following way (Haldane and Priestley).* A piece of indiarubber tubing is taken, of about 1 inch diameter and 4 feet long. Into one end (Fig. 490) is fitted a mouthpiece, near to which is a gas sampling-tube, which is provided with three-way taps at both ends. Before an experiment the sampling-tube is completely evacuated. The subject of the experiment, after breathing *normally* a few times, puts his mouth to the tube at the end of a normal inspiration, expires quickly and deeply, and closes the mouthpiece with his tongue. The tap of the sampling-tube is then turned, and the air near the mouthpiece, which is that last expelled from the lungs, rushes into it. The tap of the tube is then turned off, and the gas sample removed for analysis. A similar sample is then taken, in which the subject expires deeply at the end of a normal expiration. This sample will contain slightly more $\mathrm{CO_2}$ and less $\mathrm{O_2}$ than that obtained at the end of inspiration. The mean of the two samples is taken as the average composition of the subject's alveolar air.

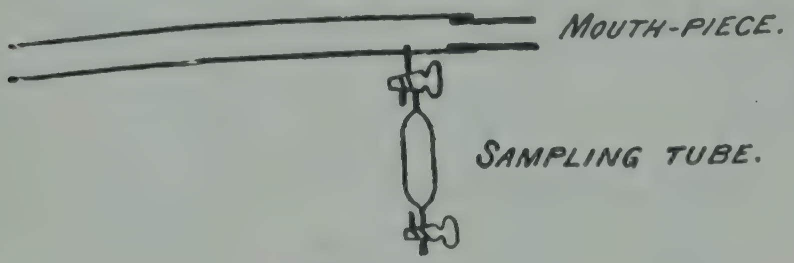


FIG. 490.

The composition of the alveolar air in a resting subject at normal atmospheric pressure is nearly constant, though showing variations from one individual to another. It also varies with posture ; the $\mathrm{CO_2}$ content may be 0·3 to 0·4 per cent. higher when supine than when standing erect, apparently because there is less blood and more air in the lungs when standing.† We may take the average composition for adult men at 760 mm. pressure, to be as follows :—

MEAN COMPOSITION OF ALVEOLAR AIR. (Adult Men.)

	Volumes per cent. (dry	Pressure mm. Hg (moist)
Carbon dioxide	5·50	39
Oxygen	13·80	99
Nitrogen	80·70	577
Water vapour	0·00	45
	100·00	760

It is often better to express the data regarding respiratory airs in terms of partial pressures, as in the last column of this table, than in terms of percentage composition. From Dalton's law we know that the total pressure exerted by a mixture of gases is equal to the sum of the separate pressures which each gas would exert if it alone occupied the whole volume. Aqueous vapour must be considered as one of the gases in alveolar air, and its partial pressure at 37° C. (= 45 mm.) allowed as a part of the total (*i.e.* barometric) pressure. The following example will explain the calculation :—

---

* HALDANE and PRIESTLEY. *J. Physiol.*, 1904, **32**, 225.

† MACKAY. *J. Physiol.*, 1943, **102**, 228.

Suppose that a sample of alveolar air had the composition (dry), CO_2 = 5·6 per cent. ; O_2 = 13·7 per cent. ; N_2 = 80·7 per cent. by volume, and that the barometer was at the time at 755 mm. Hg.

Then, since aqueous vapour = 45 mm. Hg, the remaining gases together contribute 710 mm. Hg, and of this

$$CO_2 = \frac{5\cdot 6}{100} \times 710 = 39 \text{ mm. Hg.}$$

$$O_2 = \frac{13\cdot 7}{100} \times 710 = 98 \text{ mm. Hg.}$$

$$N_2 = \text{the remainder (or } \frac{80\cdot 7}{100} \times 710) = 573 \text{ mm. Hg.}$$

These are the pressures exerted by each gas on any surface boundary.

The range for CO_2 pressures is about 32 to 45 mm. for men. In children and women the pressure of carbon dioxide is on the average a little less, and the oxygen a little more.

The difference between the composition of expired air and alveolar air is determined by the dilution of the alveolar air with that contained in the dead space. Hence, with shallow breathing there will be a large difference, but this will decrease with increased depth of respiration. Thus, if the alveolar air contained 6 per cent. CO_2 and the dead space amounted to 150 c.c., the expired air would contain only 3 per cent. CO_2 when the person was taking in only 300 c.c. at each respiration. If, however, he was breathing slowly and deeply, so as to raise the tidal air to 1,500 c.c., only one-tenth of this would be represented by the dead space, and the expired air would contain nine-tenths as much CO_2 as the alveolar air, *i.e.* 5·4 per cent.

The dead space can be determined by measuring the volume and estimating the composition of the air in an average expiration, and by finding the composition of the alveolar air. Then dead space

$$= \frac{\text{Tidal air vol.} \times (CO_2 \text{ per cent. in alveolar air} - CO_2 \text{ per cent. in expired air})}{CO_2 \text{ per cent. in alveolar air}}$$

The changes in the composition of alveolar air with each breath are by no means so great as those produced in the tidal air, since the latter forms only a small proportion of the total air in the lung alveoli. Thus, at the end of a normal expiration, the alveoli still contain 2,500 c.c. of gases. In inspiration 360 c.c. atmospheric air are taken into this space and mixed with the 2,500 c.c. already there. The " ventilation coefficient " in quiet breathing is therefore only one-seventh, and the change in the oxygen and carbon dioxide content of the alveolar air produced by this access of 360 c.c., will amount to less than one-half per cent. This is illustrated by the following figures from Haldane and Priestley,* giving the alveolar content in carbon dioxide at the end of inspiration and at the end of expiration respectively.

ALVEOLAR CO_2 PER CENT.

Individual	Alveolar CO_2 at end of inspiration. (Mean of twelve observations) per cent.	Alveolar CO_2 at end of expiration per cent.	Mean
J. S. H.	5·54	5·70	5·62
J. G. P.	6·17	6·39	6·28

We can thus speak of an average composition of alveolar air, which, in spite of the constant ventilation, remains almost constant in composition.

---

* HALDANE and PRIESTLEY. *J. Physiol.*, 1904, **32**, 222.

THE RESPIRATORY FUNCTION OF THE BLOOD

The Blood Gases. It was shown by Magnus, in 1837,* that the blood passing to the lungs contained more carbon dioxide and less oxygen than that passing away from the lungs. This proved that combustion does not occur in the lungs, as Lavoisier had suggested,† but that the blood acts simply as a carrier of the oxygen from the lungs to the tissues, and of the carbon dioxide from the tissues to the lungs. We thus learnt to distinguish between external and internal respiratory processes. If any further proof is required that it is the tissues themselves which utilise oxygen and produce carbon dioxide, we may refer to the fact that any surviving tissue, quite free from blood, can be shown by the use of the microrespirometer (p. 146), to absorb oxygen and give out carbon dioxide. A consideration of the process of external respiration necessitates an investigation of the manner in which gases are held by the blood and of the factors which are responsible for the transfer of oxygen and carbon dioxide from alveolar air to blood, and from blood to alveolar air.

If blood be exposed to a vacuum, preferably at body temperature, the whole of its contained gases are given off. In this manner, about 60—70 c.c. of mixed gases may be obtained from 100 c.c. of blood; they consist of oxygen, carbon dioxide and nitrogen (with argon in insignificant quantities). The nitrogen forms only about 1 volume per cent. in both arterial and venous blood. The amounts of oxygen and carbon dioxide in these two kinds of blood differ, however, within wide limits. The following Table ‡ represents the average composition of the gases obtained from arterial and venous bloods in man:

From 100 vols.	May be obtained		
	Of oxygen	Of carbon dioxide	Of nitrogen
Of arterial blood .	19·0 ± 0·49 vols. .	49·9 ± 0·51 vols. .	0·9 vols.
Of venous blood .	10·6 ± 0·36 ,, .	57·8 ± 0·35 ,, .	,, ,,

Measured at 760 mm. and 0° C.

Simple evacuation, though a standard method, is very laborious, and demands relatively large amounts of blood. The modern methods (p. 498), enable accurate determinations to be made on 0·5 to 2·0 c.c. of blood.

Laws of Solubility of Gases in Liquids. If a gas is in simple solution, the amount of it dissolved varies directly with its partial pressure. When only one gas is present, the whole pressure is due to that one gas. Thus, if water dissolves a certain quantity of a gas at a given temperature and pressure, it will take up twice as much if the pressure of the gas be doubled. The absorption coefficient of a liquid for a gas at a given temperature is the number of cubic centimetres of gas, measured at N.T.P., which will be taken up at the particular temperature by 1 c.c. of the liquid when the gas is at a pressure of 760 mm. Hg. The absorption coefficient diminishes with rise of temperature. The following table represents the absorption coefficients for some important gases in water at various temperatures:

Temperature	Oxygen	Carbon dioxide	Carbon monoxide	Nitrogen §
0	0·0489	1·713	0·0354	0·0239
10	0·0380	1·194	0·0282	0·0196
20	0·0310	0·878	0·0232	0·0164
30	0·0262	0·665	0·0200	0·0138
40	0·0231	0·530	0·0178	0·0118

From this table we see that 100 c.c. of water at 0° C. will absorb 4·89 c.c oxygen at one atmosphere pressure. If the pressure is reduced to one half a atmosphere, the volume of gas absorbed, when measured at N.T.P., will be 2·45 volumes. On a graph, of which the ordinates represent the amount of gas dissolved, and the abscissæ the different pressures of the gas, we shall get a straight line. The amount absorbed depends on the partial pressure of the particular gas considered, and is not altered by the presence of other gases at the same time. The normal pressure of the whole atmosphere is 760 mm. Since the atmosphere consists roughly of four parts of nitrogen with one part of oxygen, the atmospheric pressure is due, as to one-fifth to the oxygen, and as to four-fifths to the nitrogen. If we shake up water at 0° C. with the atmospheric air at the ordinary pressure, 100 c.c. of water will absorb 4·89 c.c. × $\frac{1}{5}$ of oxygen, and of nitrogen 2·39 c.c. × $\frac{4}{5}$. We may therefore extend our statement as to the solubility of gases in liquids, and say that *the amount of gas which a liquid will dissolve is directly proportional to the partial pressure of the gas* (Henry's law).

---

* Magnus. *Ann. d. Physik.*, 1837, **40**, 583; 1845, **66**, 177.

† Seguin and Lavoisier. *Mém. Acad. Sciences*, 1790, 77.

‡ After data from Looney and Jellinek. *Am. J. Physiol.*, 1937, **118**, 225.

§ Atmospheric "nitrogen" contains 1·185 per cent. argon, and this mixture is slightl less soluble than pure nitrogen, *e.g.* 0·0134 at 30° C.

When water is saturated with a gas at a given pressure, a state of equilibrium exists between the gas dissolved and the gas in contact with the liquid. At equilibrium, the number of molecules of the gas entering the liquid in unit time must equal the number of molecules of the gas leaving it. If we remove the liquid after saturation, say, at one atmosphere, to a vessel where it is in contact with the same gas at a pressure of half an atmosphere, the liquid will give off gas until the amount left in solution is diminished to one-half. The gas dissolved in a liquid thus exerts a pressure or *tension* tending to make it escape from the liquid. The only way in which we can measure this tension is by finding what pressure of that gas is in equilibrium with the liquid. Thus, if we take some water containing carbon dioxide in solution, divide it into two parts, and shake up one part with a gaseous mixture containing 4 per cent. of carbon dioxide and the other part with a mixture containing 5 per cent. of carbon dioxide, and find that the solution loses gas to the former and takes up carbon dioxide from the latter, we may conclude that the tension of carbon dioxide in the original liquid was something between 4 and 5 per cent. of an atmosphere. It is by similar means that the tensions of gases in the blood are measured, the instruments for this purpose receiving the name of *aerotonometers*.

The solvent power of water for gases is diminished if the water contains solid substances in solution. The absorption coefficient of blood plasma for gases is 97·5 per cent. of that for pure water, and of blood 92 per cent., that of the blood corpuscles being as low as 81 per cent. owing to their smaller content of free water. We may thus compute the absorption coefficient (for *physical* solution) of blood plasma, blood and blood corpuscles for oxygen, nitrogen and carbon dioxide.

	Oxygen		Nitrogen		Carbon dioxide	
	15°	38°	15°	38°	15°	38°
Blood plasma . .	0·033	0·023	0·017	0·012	0·994	0·541
Blood	0·031	0·022	0·016	0·011	0·937	0·511
Blood corpuscles .	0·025	0·019	0·014	0·010	0·825	0·450

THE CARRIAGE OF OXYGEN BY THE BLOOD

The Oxygen Dissociation Curve. From the above Table we see that 100 volumes of blood at 38° C. would contain 2·2 vols. of oxygen *in solution*, if the blood had been brought into equilibrium with oxygen at a pressure of one atmosphere. The blood in the lungs is, however, exposed to air which contains only about one-sixth of its volume of oxygen, so that the total amount of oxygen present in physical solution in arterial blood cannot be more than about 0·36 vols.

---

[*] Van Slyke and his collaborators have estimated the solubility co-efficient of CO_2 in blood plasma at 38° C. as 0·510 instead of 0·541 as given by Bohr.

per cent. But, as a matter of fact, blood saturated with air contains about twenty volumes per cent. of oxygen, so that most of the oxygen cannot be in merely physical solution, but must be in some form of reversible combination with some of the constituents of the blood ; and, in fact, practically the whole is contained in the red blood corpuscles, in combination with hæmoglobin, the plasma containing no more than can be accounted for by simple solution. This is further proved by the fact that we can displace the oxygen (O_2) from the oxyhæmoglobin by equal volumes of CO or NO.

The relation between the partial pressure of oxygen and the proportion of oxyhæmoglobin present (or "percentage saturation") when equilibrium is reached under varying conditions, can be investigated in the following way :

A cylindrical glass vessel of about 300 c.c. capacity is filled with a gaseous mixture of known composition, containing oxygen. Into it are introduced 2 or 3 c.c. of blood or of hæmoglobin solution. It is then immersed in a horizontal position in a water bath at a constant temperature, and slowly revolved. In this way the blood is continually spread in a thin layer over the sides of the vessel. At the end of half an hour it will have attained equilibrium with the gaseous mixture, and a sample may be drawn off for determination of its percentage oxygen saturation. A sample of the gas is also analysed to find the final partial pressure of its components. The results, plotted as in Fig. 491, where abscissæ are pressures of oxygen in mm. Hg, and ordinates are percentage saturations, form what is called the "*dissociation curve*" of the blood (or of the oxyhæmoglobin). It tells us what percentage of the hæmoglobin is saturated with oxygen at each partial pressure of that gas.

It will be noted that, as the partial pressure of oxygen is lowered, little oxygen is given off until the partial pressure of the oxygen is diminished to about 30 mm. Hg (Fig. 491). At this point a rapid evolution of oxygen begins, and continues at falling pressure until, at 0 mm. pressure, all the oxyhæmoglobin is converted into hæmoglobin. The same observation may be made in the reverse direction, *i.e.* if a solution of reduced hæmoglobin be exposed to gradually increasing pressures of oxygen, it will be found that the greatest absorption takes place between 0 and 30 mm. Hg. After this point, the oxygen is more and more slowly absorbed until it asymptotically reaches the point of complete saturation.

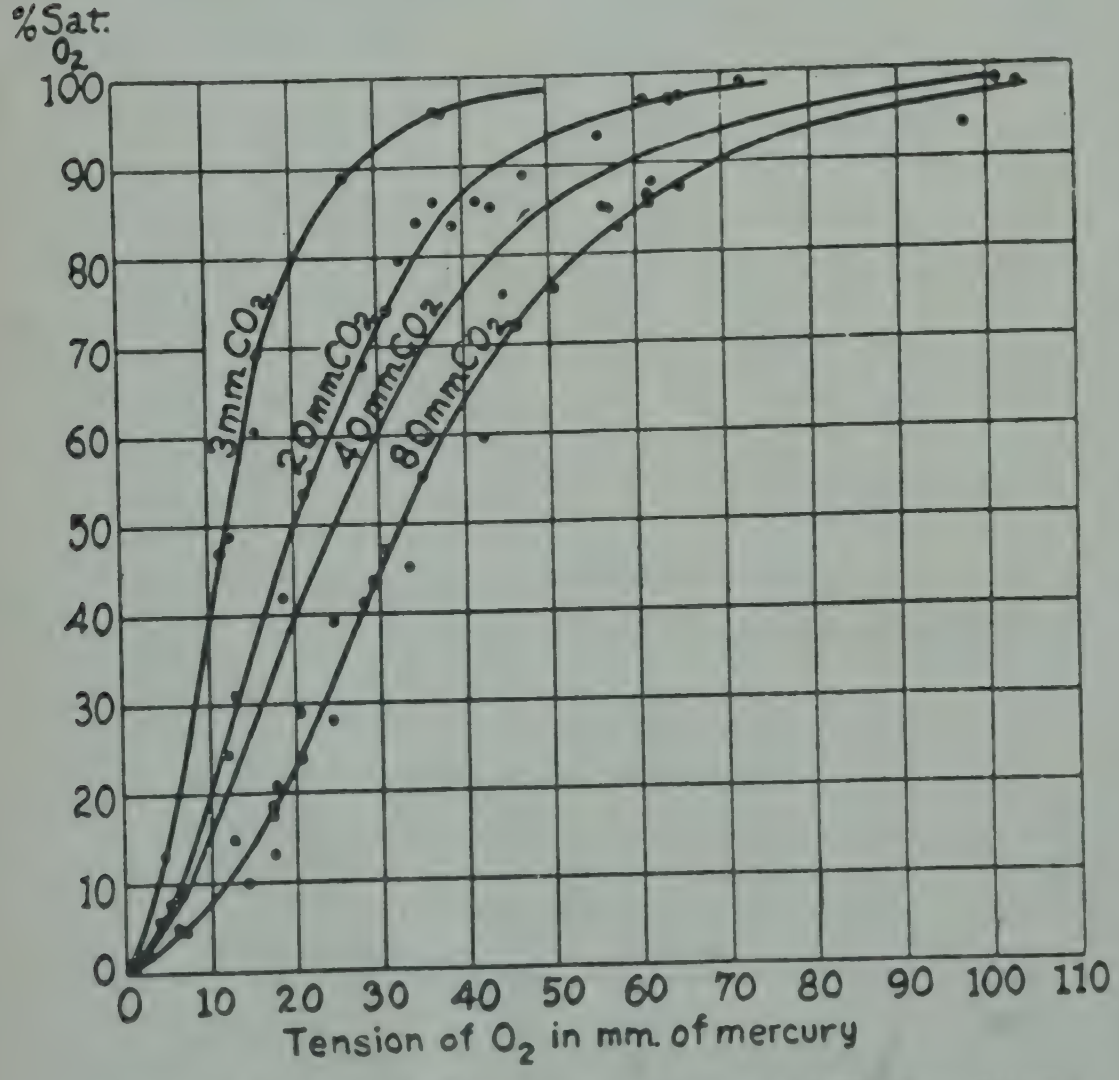


FIG. 491. Oxygen Dissociation Curves of Human Blood, exposed to pressure of 3, 20, 40 and 80 mm. of CO_2. (BOCK, FIELD and ADAIR. *J. Biol. Chem.*, 1924, **59,** 366.)

The dissociation curve approximately fits the equation (Hill's equation [*]) :—

$$\frac{y}{100} = \frac{Kx^{2.5}}{1 + Kx^{2.5}}$$

---

[*] A. V. HILL. *J. Physiol.*, 1910, **40,** *Proc.* iv. ; W. E. L. BROWN and A. V. HILL. *Proc. R. S.*, 1923, **94B,** 297.

where x = oxygen pressure, y = per cent. saturation of the Hb, and K the dissociation constant. Hence if one point is actually determined, the rest of the curve can be approximately calculated. The theoretical significance of the curve is doubtful.* The fact that the curve is not a rectangular hyperbola, though somewhat similar to one, is important. The S-shaped part which it shows at the lower oxygen pressures means that with falling oxygen pressures the hæmoglobin gives up its oxygen more easily than the rectangular hyperbola would require. It is stated to be due to the formation of intermediary compounds of hæmoglobin with oxygen.

The dissociation curve of oxyhæmoglobin is largely altered by slight variations in the fluid in which the hæmoglobin is dissolved. The most important of these conditions are (1) the saline content of the fluid (as, *e.g.* in blood), (2) the reaction of the fluid, (3) the source of the hæmoglobin, *i.e.* the age and species of the animal from which it was prepared.

The effect of change of hydrogen-ion concentration, or so-called "Bohr effect," is important. In Fig. 491 is represented the influence of varying tensions of carbon dioxide on the dissociation curve. It will be seen that the higher the CO_2 tension in the blood, the less firmly is its oxygen held. In other words, carbon dioxide, or other acid, tends to expel oxygen from the blood. The normal tension of CO_2 in arterial blood is about 40 mm. Hg. At 150 mm. oxygen pressure the blood is practically saturated with oxygen, whatever (within physiological limits) the pressure of the carbon dioxide; but at lower pressures of oxygen, the pressure of carbon dioxide makes a considerable difference to the oxygen saturation. Thus, at an oxygen pressure of 20 mm. Hg the amount of oxyhæmoglobin formed is 80 per cent. when the carbon dioxide pressure is 3 mm., whereas at a pressure of carbon dioxide of 40 mm. the degree of saturation of the hæmoglobin is only 40 per cent. In consequence of this fact, in the tissues where the carbon dioxide tension is high, the oxyhæmoglobin will be dissociated with greater ease, so that oxygen will readily be set free where it is most wanted.

A change in the opposite direction is produced when part of the hæmoglobin is converted into carboxy-hæmoglobin or methæmoglobin,† the curve of dissociation being moved to the left and having its S-inflection reduced, *i.e.* the oxygen is more difficult to remove from the hæmoglobin.

Another important factor which affects the oxyhæmoglobin dissociation curve is, of course, the temperature at which equilibrium is effected. Rise of temperature, like rise of H-ion concentration, flattens out the curve, *i.e.* rise of temperature facilitates the removal of oxygen from oxyhæmoglobin.‡

Uptake of Oxygen by the Blood. The foregoing considerations relate to the state of equilibrium between hæmoglobin and oxygen, but give no information as to the speed with which that state is reached. In passing from alveolar air to oxyhæmoglobin, the following phases must be considered in the transport of oxygen, and all must take time :—

1. Passage through the alveolar membrane.
2. Passage into the blood plasma.
3. Passage through the erythrocyte envelope.
4. Diffusion into the erythrocyte cytoplasm.
5. Combination with hæmoglobin.

All phases together are expressed in the so-called diffusion constant (*v.* p. 740), but here we are concerned with phases 2 to 5.

Phase 5, the velocity of the reaction $Hb + O_2 \rightarrow HbO_2$ is so high ($k = 132$) that it may be regarded for physiological purposes as instantaneous : the speed of the reverse reaction of reduction, though less, is also high ($k = 17{\cdot}5$).§

An idea of the velocities concerned in phases 2 to 5 collectively can be obtained by measuring the speed of uptake (and conversely, of the yielding-up) of oxygen by the blood. The velocity of the exchanges is probably about one-tenth of that for phase 5 alone, but will clearly depend on the extent of surface of blood exposed to the gas phase, and if this is not large the speed will be still less. Fig. 492 illustrates these velocities under comparable, but artificial, conditions for uptake and release of oxygen (and of CO_2): in the better conditions in the body these speeds would no doubt be considerably greater.

---

* McLean. *Physiol. Rev.*, 1938, **18,** 495.

† Darling and Roughton. *Amer. J. Physiol.*, 1942, **137,** 56. Roughton and Darling. *Ibid.*, 1944, **141,** 17.

‡ Barcroft and King. *J. Physiol.*, 1909, **39,** 374.

§ Hartridge and Roughton. *Proc. R. S.*, 1925, **107A,** 654 ; *J. Physiol.*, 1926, **62,** 232.

Uptake of Oxygen in the Lungs. We are now in a position to understand how the oxygen is taken up by the blood, as it circulates round the pulmonary alveoli. Arterial blood, such as that which fills the pulmonary veins and the systemic arteries, is very nearly (viz. about 98 per cent.) saturated with oxygen (*e.g.* it contains about 19·5 volumes out of a possible 20 volumes per cent.). Venous blood, as passed to the lung, is about 66 per cent. saturated (*i.e.* it contains about 13 volumes of oxygen per cent.). The oxygen tensions are about 97 mm. for arterial blood, and 40 mm. for venous blood, *i.e.* at a tension of 97 mm. the blood is nearly saturated. The pressure of oxygen in the alveoli is normally about the same, being usually about 98 mm. Hence the venous blood brought to the alveoli, being only separated from the alveolar air by the thin alveolar membrane, and, aided by the Bohr effect caused by the release of CO_2 from the blood, will rapidly take up oxygen from it, and will speedily approach the point of saturation. In its rapid transit through the lungs, however, final equilibrium may not be reached, because, as shown in Fig. 492, the rate of uptake of oxygen slows down as saturation is neared. Also, the arterial blood is never quite in equilibrium with alveolar air, because some venous blood from bronchial veins drains into the pulmonary veins. (In some pathological cases also, part of the pulmonary circulation leads through collapsed or otherwise non-ventilated alveoli. In such cases even breathing pure oxygen fails to give full saturation of the arterial blood.)

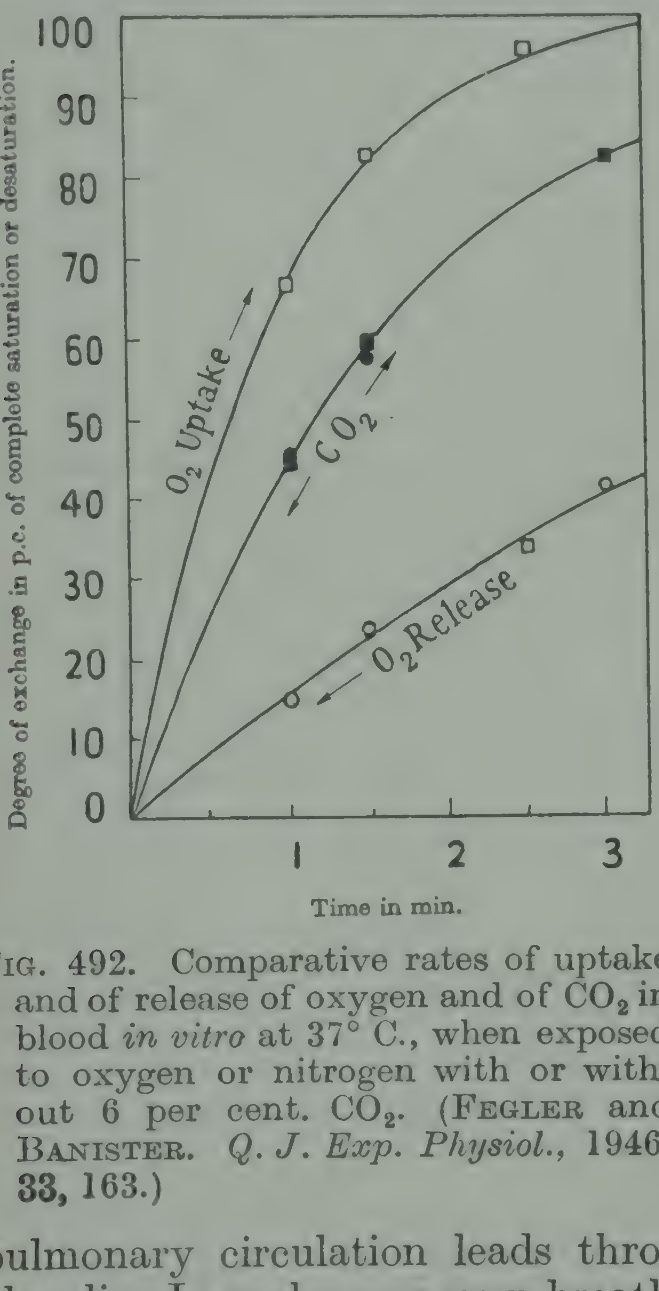


FIG. 492. Comparative rates of uptake and of release of oxygen and of CO_2 in blood *in vitro* at 37° C., when exposed to oxygen or nitrogen with or without 6 per cent. CO_2. (FEGLER and BANISTER. *Q. J. Exp. Physiol.*, 1946, **33,** 163.)

Passage of Oxygen to Tissues. The blood, thus laden with oxygen, is sent from the left side of the heart to all parts of the body. It must be remembered that in the tissues, as in the lungs, the hæmoglobin does not exchange its oxygen directly. In both cases, the interchange is effected through the intermediation of capillary wall and of the plasma and, in the tissues, of the tissue fluid as well. Since the tissue elements are constantly using up oxygen, they absorb any oxygen that is present in the surrounding fluid. There is, in consequence, a descending slope of oxygen tensions from red blood corpuscle through plasma, vessel wall, tissue fluid and tissue element and (assuming that all the media are permeable to it) the oxygen is bound to follow this gradient. The cell draws from the tissue fluid, and the tissue fluid from the plasma, so that the oxygen tension in the plasma sinks. The immediate result is an evolution of oxygen from the oxyhæmoglobin and the oxygen set free is taken up by the plasma, to be in turn passed on to the tissue fluid and cells.

THE CARRIAGE OF CARBON DIOXIDE BY THE BLOOD

Carbon dioxide is soluble in water, and therefore in blood, which at 38° C., when exposed to a pressure of 40 mm. CO_2, *i.e.* about that of arterial blood, would hold, in a state of *physical* solution, about 2·7 volumes of CO_2 per 100 c.c. of blood. But blood, as drawn from the body, may contain as much as 60 volumes per cent. of carbon dioxide, held in such a state as to be removable in a vacuum, and we have now to inquire why it is that blood is capable of holding such large amounts of carbon dioxide. A further fact is that this carbon dioxide can also be expelled from it by the addition of strong acids. This suggests at once that the carbon dioxide is held in the blood as a bicarbonate or carbonate, from which union it is expelled by acid, or when the blood is exposed to a vacuum. It should further be noted that, after all possible carbon dioxide has been drawn from blood by exposure to a vacuum, no more can be driven off on addition of a strong acid. One method of expulsion is as good as the other, so far as the total yield of CO_2 from whole blood is concerned.

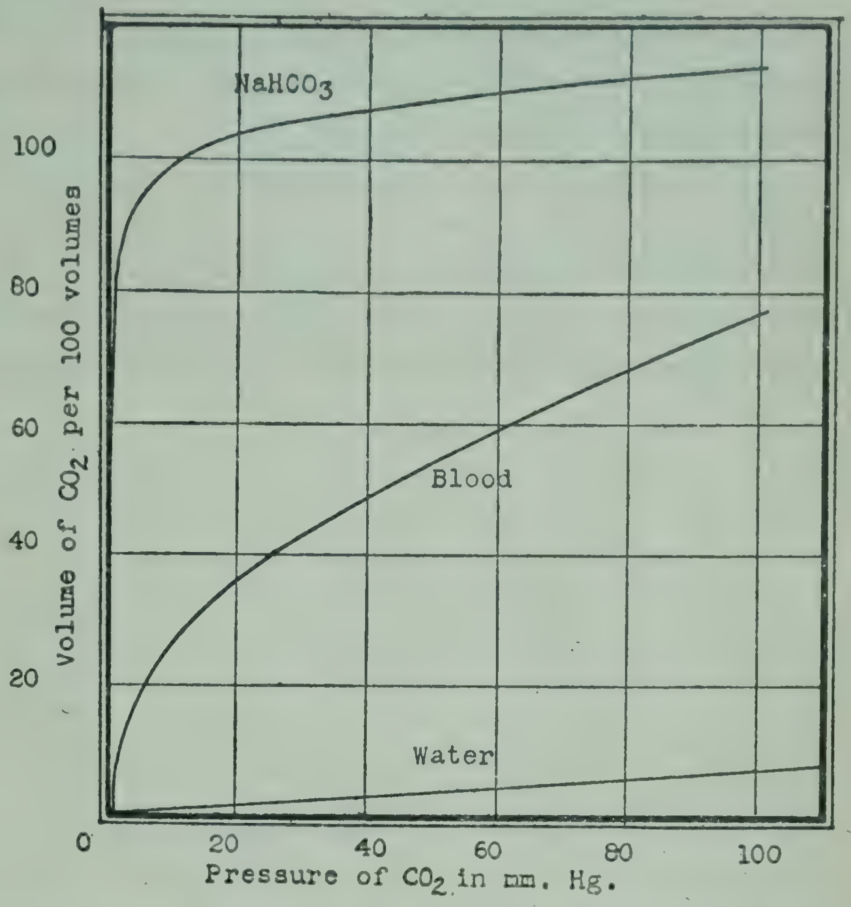


FIG. 493. Comparison between the CO_2 Dissociation Curves of Blood and of Sodium Bicarbonate Solution of a concentration of 0·0484 N. The amounts of CO_2 held in physical solution in water are also shown (lowest line). (Redrawn from data by PARSONS. *Journ. Physiol.*, 1919, **53**, 42, 340.)

Carriage as Bicarbonate. Valuable information has been obtained by the study of the CO_2 dissociation curves of blood, and of somewhat similar chemical systems. These are made by saturating samples of the fluid with a series of gas mixtures of varying CO_2 pressure, and then determining the volume of CO_2 which can be expelled from the fluid by addition of a strong acid.

The CO_2 dissociation curve of blood when plotted on logarithmic paper gives a straight line, the slope of which depends on the oxygen capacity.* The slope between 60 mm. and 30 mm. CO_2 can be calculated from the equation :—

$$\Delta\,(CO_2)_{60-30} = 0{\cdot}334\,(O_2\ \text{Cap}) + 6{\cdot}3.$$

Using this, a curve can be constructed from one point.

If we compare the CO_2 dissociation curve of blood with that of a solution of sodium bicarbonate, containing the same amount of available base (0·0484 N.), it is evident, as is shown in Fig. 493, that there is a very great difference between them. For, whereas all the carbon dioxide is removed from the blood when the pressure of CO_2 is reduced to zero (*i.e.* in a vacuum), it is impossible, by lowering the pressure, to remove from the bicarbonate solution more carbon dioxide than corresponds to its conversion to carbonate, according to the equation—

$$2NaHCO_3 = Na_2CO_3 + CO_2 + H_2O\,;$$

i.e. from an aqueous bicarbonate solution only half as much carbon dioxide can be expelled by a vacuum as by addition of a strong acid :—

$$2NaHCO_3 + 2HCl = 2NaCl + 2CO_2 + 2H_2O.$$

The carbon dioxide in blood is chiefly bound in the plasma in the form of sodium bicarbonate. This is shown by the fact that the amount of carbon dioxide which can be expelled from the plasma by addition of a strong acid agrees with the amount of base found by titration, and further, all this bicarbonate can be removed as such by dialysis. If plasma is saturated with carbon dioxide at atmospheric pressure, practically all the available base of the plasma will be present in the form of sodium bicarbonate. The amount of bicarbonate present in the plasma under any circumstances can readily be determined, and is called the "*alkali reserve.*" *

---

* PETERS *et al.* *J. Biol. Chem.*, 1923, **56**, 745 ; **58**, 769.

If we examine the dissociation curve of blood plasma, we find that separated plasma behaves less like blood than like an aqueous bicarbonate solution (Fig. 494). This is confirmed by the well-known fact that, although all the CO_2 can be removed from blood by a vacuum, this cannot be done with blood plasma; after plasma has been thoroughly pumped out *in vacuo*, addition of acid can still set free carbon dioxide from it. It is clear then, that if the carbon dioxide *is* fixed in the form of bicarbonates in the blood, there must also be in the blood some substance which, while not interfering with the taking up of carbon dioxide by the base when the blood is exposed to increasing pressures of carbon dioxide, will yet act as an acid does, and drive out carbon dioxide, ultimately to completion, as the partial pressure of CO_2 falls. *This substance must be contained in the corpuscles.* It is the corpuscles which *in vacuo* act like an acid, expelling carbon dioxide. This can be proved by the simple experiment of mixing thoroughly well-evacuated plasma with corpuscles, when CO_2 is at once evolved on further evacuation.

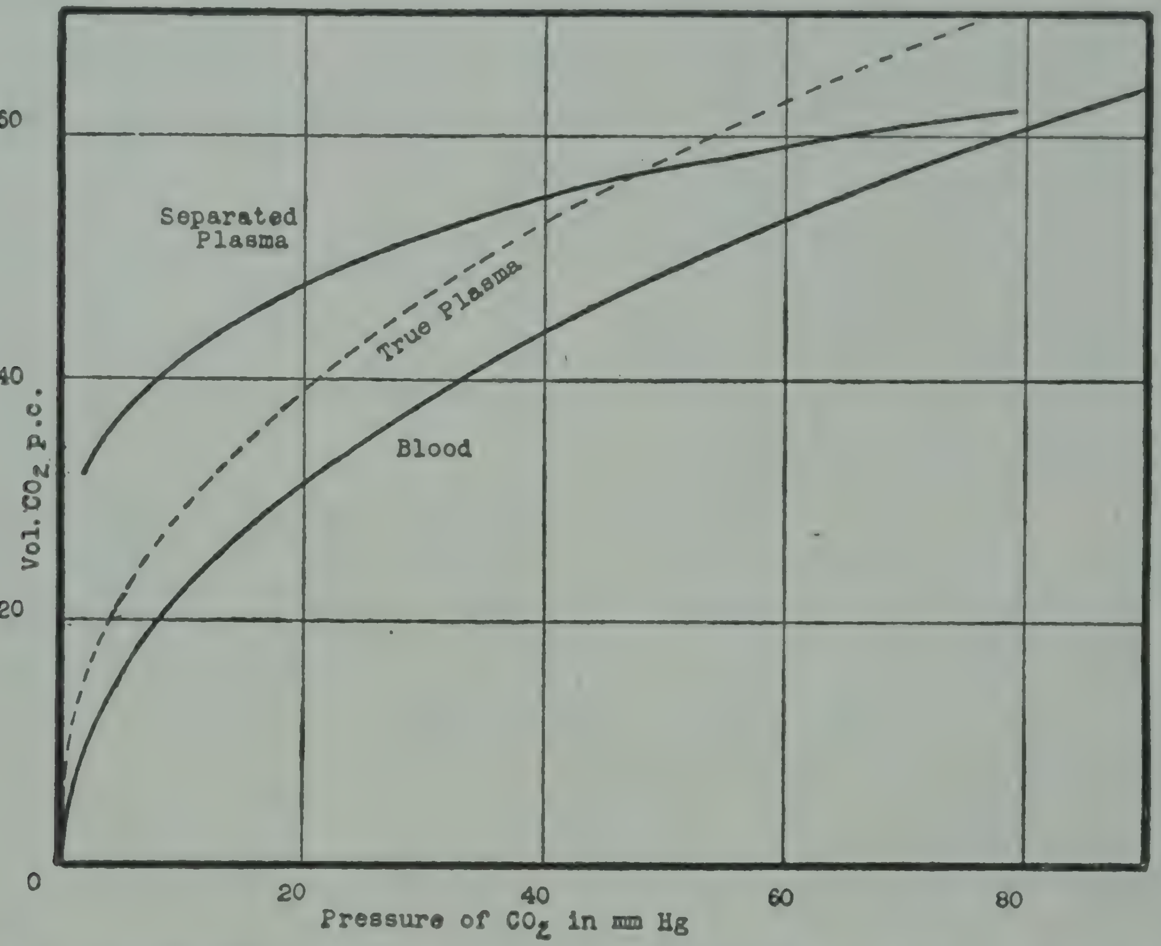


FIG. 494. Comparison between the CO_2 Dissociation Curves of Blood, and of Separated Plasma and true Plasma from the same Blood. (Redrawn from data by JOFFE and POULTON. *J. Physiol.*, 1920, **54,** 129.)

It is also shown by the comparison of true plasma with separated plasma. The name "*true plasma*" is given to plasma drawn direct from the blood, without any change in its gas content. Specimens of true plasma are obtained by centrifuging the blood under a layer of liquid paraffin, to hinder exchange of gases at the surface, and then pipetting off the plasma layer. This represents the plasma which was in a state of equilibrium with the corpuscles of the blood when this was centrifuged.

Now, if blood be exposed to gas mixtures containing different pressures of carbon dioxide, as in making a CO_2 dissociation curve, it is a simple matter to separate off samples of true plasma corresponding to each pressure of CO_2, to determine their bicarbonate content, and thus to obtain a curve which may be called the dissociation curve of true plasma. Such a curve represents the carbon dioxide content of plasma which at each point is in equilibrium with the corpuscles. When this is done, it at once becomes evident that the true plasma behaves in a way entirely different from the separated plasma discussed above. Fig. 494 illustrates this difference.

---

* Strictly speaking, the "alkali reserve" is defined as the volume per cent. of CO_2 yielded by the plasma on treatment with acid after the plasma has been previously brought into equilibrium with a gas mixture containing CO_2 at 40 mm. pressure (alveolar air).

The true plasma shows a curve which has much the same slope as that for blood, but at a higher level; the plasma of blood, when in equilibrium with its corpuscles, has sometimes a higher percentage bicarbonate content not only than separated plasma, but even than the whole blood. This is because the carbon dioxide-fixing power of corpuscles is less than that of plasma.

The explanation of the way in which the corpuscles confer upon the plasma the power of fixing CO_2 by the formation of bicarbonate was given by Hamburger,* who showed that when carbon dioxide is added to blood the following changes occur :—

(1) The bicarbonate content of the plasma is increased.
(2) The chloride content of the plasma is diminished.
(3) The chloride content of the corpuscles is increased.

The cations (base) of the plasma and corpuscles are unchanged.

The reason why blood can carry much more carbon dioxide than can water is that the H· ions formed in the reactions :

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^{\cdot} + HCO_3{}'$$

are all mopped up by buffers in the blood, so that the reaction is accelerated from left to right. In consequence of this the CO_2 carrying power of the blood runs parallel to its buffering power. Some of the buffers of the blood are present in the corpuscles and some in the plasma : in the corpuscles we have the non-diffusible anions of hæmoglobin, diphosphoglycerate and adenosine triphosphate, while in plasma we have small amounts of phosphate, and the plasma proteins.

Since the red cells are freely permeable to CO_2, and since hæmoglobin is quantitatively the most important blood buffer, it follows that when carbon dioxide enters the blood and reacts with potassium hæmoglobinate, far more $HCO_3{}'$ ions are formed in the corpuscles than in the plasma. The red cells are also easily permeable to $HCO_3{}'$ ions, and so, since there is a gradient of $HCO_3{}'$ from cells to plasma, this ion will tend to migrate out again into the plasma. But the envelope of the red cell is not permeable to cations, and so the $HCO_3{}'$ ions cannot be accompanied out by the positively charged ions, such as $K^{\cdot}$. Hence, owing to the tendency of the $HCO_3{}'$ to migrate, the cell becomes positively charged, and this, so to speak, sucks into it negatively charged ions, such as Cl′, which are abundant in plasma. This offers an explanation of the so-called Hamburger interchange, viz. that as CO_2 is added to blood the plasma bicarbonate increases and the plasma chloride diminishes. The advantage of this is that it enables the excellent buffering power of the corpuscles to be shared out with the poor buffering power of the plasma.

The buffering functions of the organic phosphates are of a different nature. Unlike hæmoglobin, their amounts in the corpuscle can change. When CO_2 or other acid is added, for instance, diphosphoglycerate is diminished, and so base is set free.

The Effect of Oxygenation on Carbon Dioxide Carriage. It was found by Christiansen, Douglas and Haldane † that reduced blood has a greater carbon dioxide capacity than oxygenated blood (Fig. 495).

In other words, oxygen tends to displace carbon dioxide from the blood, much as carbon dioxide is known to displace oxygen. They explained this phenomenon on the hypothesis that reduced hæmoglobin is a weaker acid than oxyhæmoglobin, *i.e.* its salts mop up $H^{\cdot}$ ions more easily.

The importance of this observation lies in the fact that during its passage along the systemic capillaries, not only is carbon dioxide added to the blood, but also oxygen is withdrawn from it, and so the buffering power of the blood is raised, thus enabling more CO_2 to be taken up and *vice versâ* in passing the lungs.

---

* HAMBURGER. *Arch. f. Phys.*, 1892, 513 ; 1893, Suppl. 153 ; *Bioch. Ztschr.*, 1918, **86,** 309.

† CHRISTIANSEN, DOUGLAS and HALDANE. *J. Physiol.*, 1914, **48,** 244.

It is of interest in this connection to attempt to form an opinion of the carbon dioxide absorption curve of blood under the conditions which prevail in the body. Such a curve (for an R.Q. of about 0·8) is shown in Fig. 495, in which portions of the carbon dioxide curves for fully reduced and fully oxygenated blood are shown on a large scale. The curve AVB represents the *physiological carbon dioxide absorption or dissociation curve.* At the point A, which may be called the "*arterial point,*" since it represents the condition in arterial blood, the fully oxygenated blood at 41 mm. CO_2 contains 52·5 volumes CO_2 per cent. The point B represents the condition if the blood were fully reduced; it would then contain, at the pressure of CO_2 of 64 mm., not 61 per cent. of CO_2 (as would be the case had there been no loss of oxygen), but 66·5 per cent. Usually the blood does not nearly reach the fully reduced state in passing through the capillaries, but would in most cases stop at a point about in the position indicated by V in the figure. This corresponds to a gain of 5 volumes of CO_2, and therefore a loss (since the R.Q. is assumed to be 0·8) of about 6·3 c.c. of oxygen. The point V is thus a typical "venous point." It will be seen that in this instance at least half the increase in CO_2 content of the blood is due to the simultaneous reduction of oxygen content.

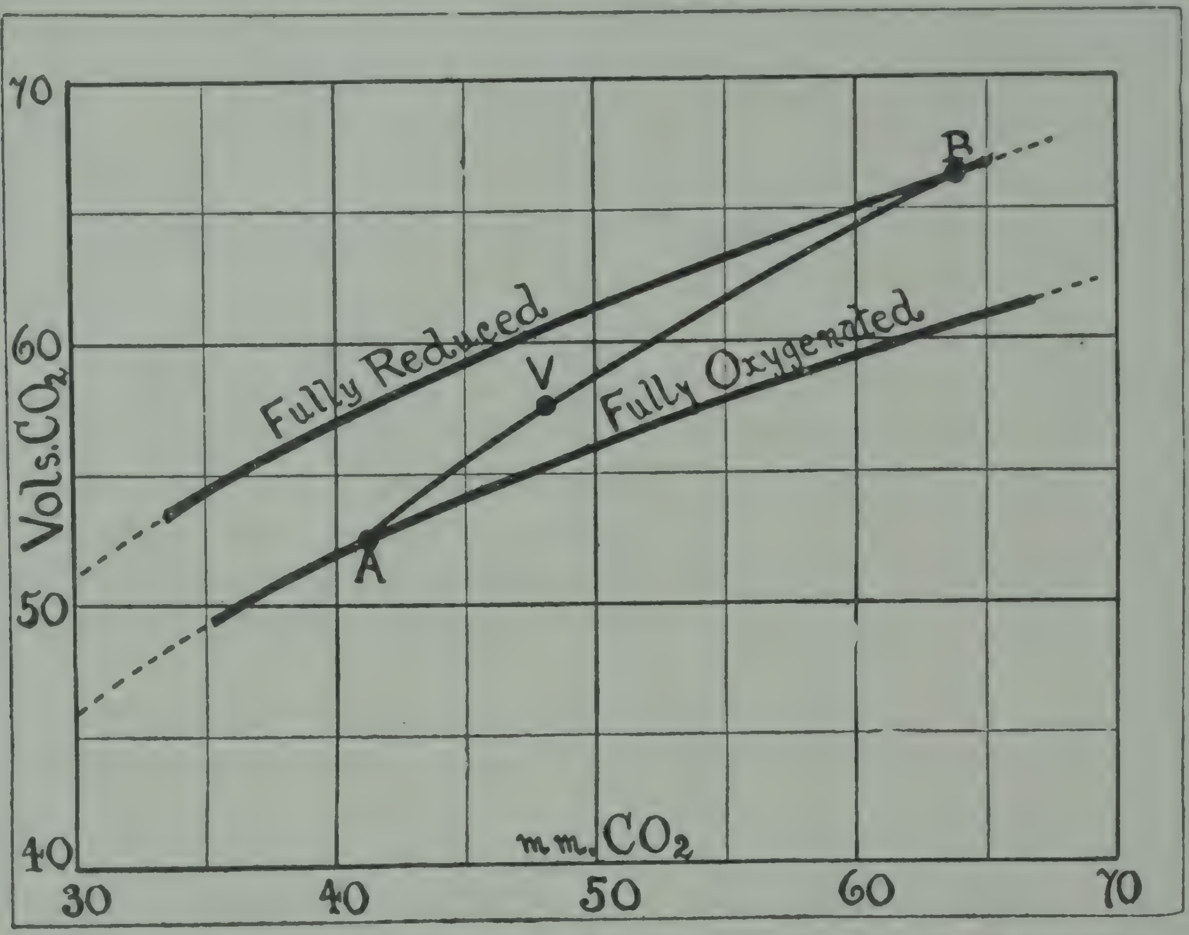


FIG. 495. CO_2 Dissociation Curves of Fully Reduced and Fully Oxygenated Human Blood (only part of the curve is shown). A V B is the physiological curve, along which the CO_2 dissociation operates in the body. (Redrawn from CHRISTIANSEN, DOUGLAS and HALDANE.)

The Blood Nomogram. We have seen that various quantities in the blood are related to more than one other quantity, *e.g.*—

(1) Oxygen saturation is related to oxygen partial pressure and to CO_2 partial pressure.
(2) CO_2 content is related to CO_2 partial pressure and to oxygen partial pressure.
(3) pH is related to (1) and (2).
(4) Plasma Cl is related to plasma CO_2, and so to (2) and (3),

and so on.

A little reflection makes it clear that all these properties of the blood are inter-related, and, in fact, L. J. Henderson showed that blood is a physico-chemical system in which many important variables are concerned. Some of these variables are :—

(1) Free oxygen (expressed by oxygen partial pressure).
(2) Combined oxygen (oxygen saturation of Hb).

(3) Free carbon dioxide (CO_2 partial pressure).
(4) Combined carbon dioxide (bicarbonate, etc.).
(5) *p*H of plasma.
(6) *p*H of corpuscles.
(7) Cl of plasma.
(8) Cl of corpuscles.
(9) Water distribution between plasma and corpuscles.

Now the relation between these and other variables for a given blood is such that, if *any two* of them are arbitrarily fixed at definite values, the remainder will thereby be fixed according to definite laws. For instance, if the p.p. of O_2 and of CO_2 are given particular values, then the *p*H of plasma and of corpuscles, the oxygen saturation, CO_2 content, Cl of plasma and of corpuscles must acquire certain values. Conversely, if we take that blood in a state of equilibrium, all the variables must have a certain value: if we alter any one of them, there will result compensatory changes in all the others when the new state of equilibrium has been reached. Henderson showed that the mutual relationships between these variables in a given blood can be expressed in the form of an alignment chart of the type shown in Fig. 496.

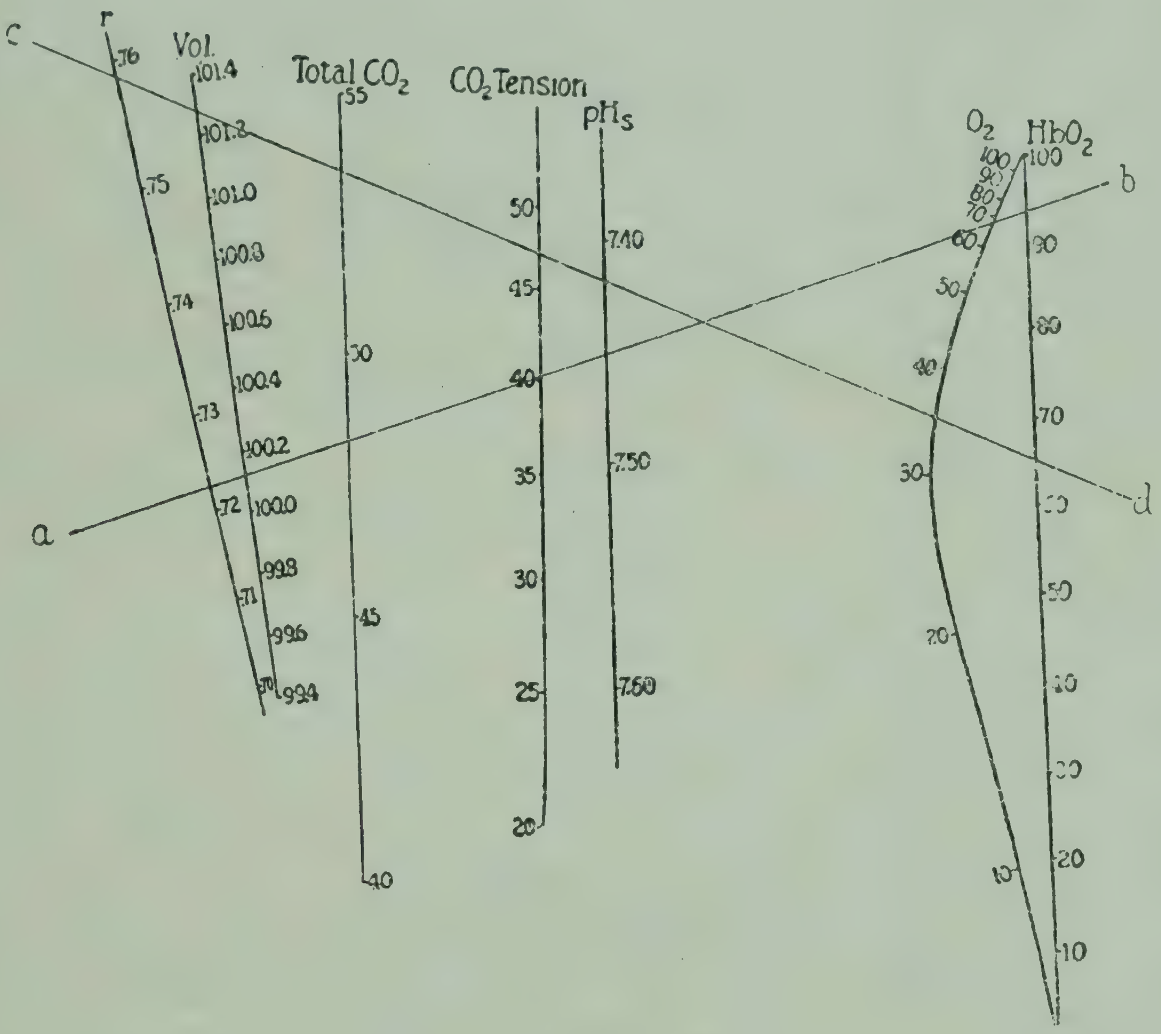


FIG. 496. Alignment Chart for a specimen of Human Blood. (From L. J. HENDERSON. *The Blood.*)

In using the nomogram of a particular individual's blood, any two determined points are joined by a straight line, which, if produced, will then cut the other points so as to give the remaining data for that blood. Thus the line *ab* might represent the conditions in arterial blood, and the line *cd* those in venous blood. Various other data are expressed in, or deducible from, this nomogram, which may be said to represent in graphic form the law of the blood. Thus, the line *r* represents the "Donnan ratio," and expresses the equilibrium between serum and corpuscles, *i.e.*,

$$r = \frac{[BHCO_3]_c}{[BHCO_3]_s} = \frac{[HCO_3]_c}{[HCO_3]_s} = \frac{[Cl']_c}{[Cl]_s} = \frac{[H]_s}{[H]_c}$$

while the line *Vol.* shows the volume of the corpuscles, as a p.c. of their volume at O_2 = 80 mm. and CO_2 = 39 mm. The line pH_s shows the *p*H of the serum (or plasma) from which, and from *r*, the *p*H of the corpuscles can be found too.

Carbonic Anhydrase. All the above facts have been found out by experiments with saturators in which blood is brought into final equilibrium with various gas mixtures—which may take from 15 to 30 minutes. The important question is, however, whether such equilibrium could occur in the short time of one or two seconds during which the blood is passing along the capillaries of lung or tissue. Most of the separate reactions are ionic ones, and therefore extremely rapid, but the reaction of formation of carbonic acid :

$$CO_2 + H_2O \rightleftharpoons H_2CO_3$$

is easily shown to be quite slow, and if uncatalysed, would, in fact, limit the whole chain of reactions to about one hundredth of the speed of those that do occur in the blood. The catalyst concerned in speeding up this reaction is an enzyme to which Roughton,* its discoverer, gave the name of *carbonic anhydrase.* Carbonic anhydrase is a protein substance which contains zinc. It accelerates the reversible reaction by which CO_2 and H_2O combine to form carbonic acid, and *vice versâ.* Its action is inhibited by cyanide (and by sulphanilamide), so that after the blood has been poisoned by addition of a trace of cyanide the reactions take place much as they do in inorganic solution.

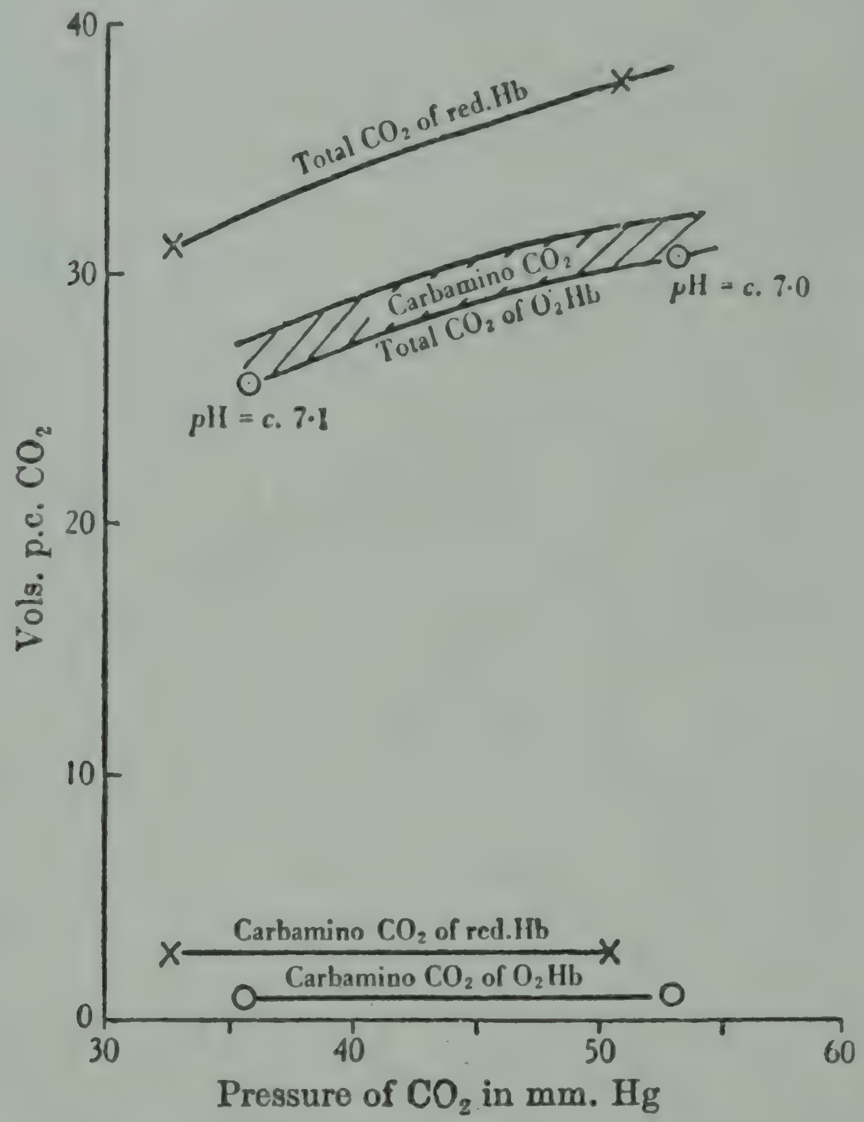


FIG. 497. Showing the CO_2 Dissociation Curves of Oxygenated and Reduced Hæmoglobin. The shaded area represents the CO_2 released from carbamino-compounds in the change from reduced to oxygenated hæmoglobin, and is seen to be a considerable fraction of the total change, although the total *amount* of carbamino-fixed CO_2, shown by the two lower lines, is never great when compared with total CO_2 content. (FERGUSON and ROUGHTON. *J. Physiol.*, 1934, **83,** 87.)

When a bicarbonate solution is acidified, we may regard its decomposition as occurring in two stages :—

$$(1)\ NaHCO_3 + HCl \rightleftharpoons H_2CO_3 + NaCl.$$
$$(2)\ H_2CO_3 \rightleftarrows CO_2 + H_2O.$$

Acceleration of reaction (2) must therefore accelerate the rate of evolution, or uptake, of CO_2. In presence of this enzyme, the rate of evolution of CO_2 from slightly acidified bicarbonate solutions is greatly accelerated. The enzyme can be precipitated by an alcohol-ether mixture, and 1 gram of the precipitate in 200 litres of solution can accelerate the evolution of CO_2 five-fold. The enzyme also accelerates the converse reaction when carbon dioxide is added to the blood. It is evident that the presence of this enzyme must considerably hasten the uptake of CO_2 in the blood in the tissues and its evolution in the lungs.

---

* BRINKMAN, MAGARIA, MELDRUM and ROUGHTON. *J. Physiol.*, 1932, **75,** Proc. 3
MELDRUM and ROUGHTON. *J. Physiol.*, 1932, **75,** Proc. 15.

Carbonic anhydrase in blood is present only in the red corpuscles. Many tissues contain traces of it, but the only other tissues which contain an appreciable amount of it are the pancreas and the stomach, in which it probably has a function connected with the secretion of bicarbonate and of HCl.

Carriage of Carbon Dioxide by Hæmoglobin. It has often been suspected that hæmoglobin may be capable of combining with CO_2, the suggestion being that the hæmatin part of the molecule combined with oxygen and the globin portion with CO_2. This anticipation has been confirmed by the work of Roughton and others.

When blood, poisoned with cyanide to eliminate the carbonic anhydrase, is exposed to CO_2 the gas is at first still taken up too quickly for the rate to be explicable on the ground that it was fixed in the form of bicarbonate, though the rate of uptake does ultimately slow down and accord with that supposition. This first quick absorption of carbon dioxide is due to the formation of *carbamino-compounds* by reaction with NH_2 groups in the protein molecule :—

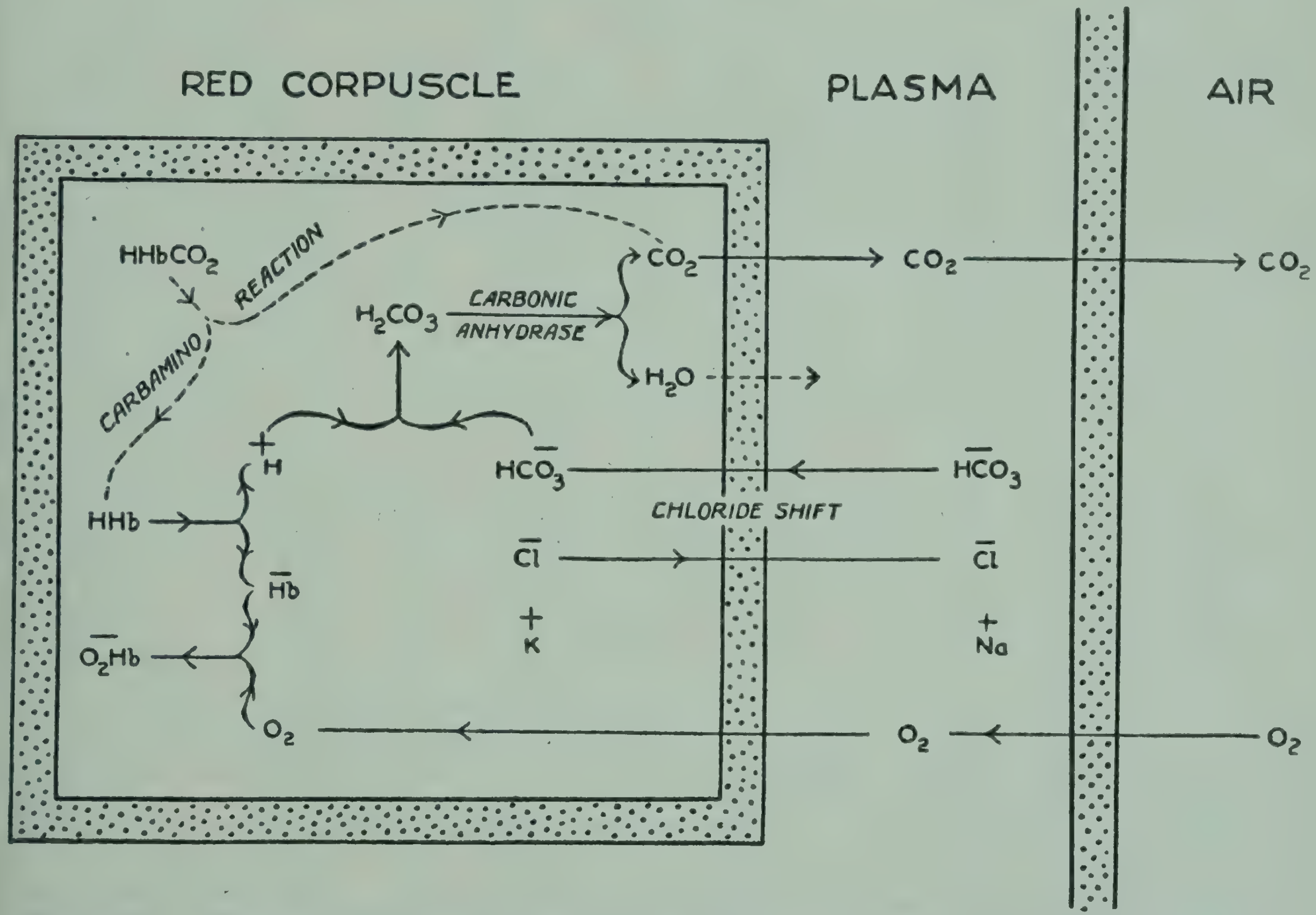


FIG. 498. Diagram of Interchanges in Corpuscle and Plasma consequent on taking up of Oxygen and expulsion of CO_2. (ROUGHTON.)

$$CO_2 + Pr.NH_2 \rightleftarrows Pr.NHCOOH.$$

The reaction is reversible, and is not interfered with by the presence of cyanides. The carbamino-compound, unlike bicarbonate or CO_2, is not precipitated by calcium or barium salts, and can actually be separated from blood in this way. In alkaline blood, and especially when the blood is deoxygenated, relatively more of the carbamino-compound is formed than when the blood is less alkaline and oxygenated. Therefore, just as with the bicarbonate CO_2, oxygenation of the blood tends to expel the carbamino-bound CO_2; although the latter may only represent 5 or 10 per cent. of the total CO_2 of the blood, it is the most mobile part, and hence is responsible for a relatively large *fraction* of the difference between the CO_2 content of oxygenated and reduced blood, as Fig. 497 shows.

The various reactions in plasma and corpuscles which take place in the lungs in connection with the uptake of oxygen by, and the expulsion of CO_2 from, the blood are shown diagrammatically in Fig. 498.

For Reference

BARCROFT (1928). The Respiratory Function of the Blood, II. Hæmoglobin. (Cambridge.)

HENDERSON, L. J. (1928). Blood. (New Haven, Yale Univ. Press and London, Oxford Univ. Press.)

ROUGHTON (1935). Recent Work on Carbon Dioxide Transport by the Blood. *Physiol. Rev.*, **15,** 241 ; (1943–4) *Harvey Lects.*, **39,** 96.

VAN SLYKE (1921). The Carbon Dioxide Carriers of the Blood. *Physiol. Rev.*, **1,** 141.

WILSON, WRIGHT (1923). Neutrality Regulations in the Body. *Physiol. Rev.*, **3,** 295.

EXCHANGE OF GASES IN THE LUNGS

A liquid gives off gas to, or takes up gas from, any other medium with which it is in contact, according to the relative pressures of the gas. The question arises whether we can account by physical means for the interchange which occurs between the alveolar air and the blood in the pulmonary capillaries. In order to answer this question we must know the partial pressures or tensions of oxygen and of carbon dioxide in the alveolar air, in the venous blood coming to the lungs, and in the arterial blood leaving the lungs. If a physical explanation is to be accepted, the partial pressures and hence the flow of oxygen must be,

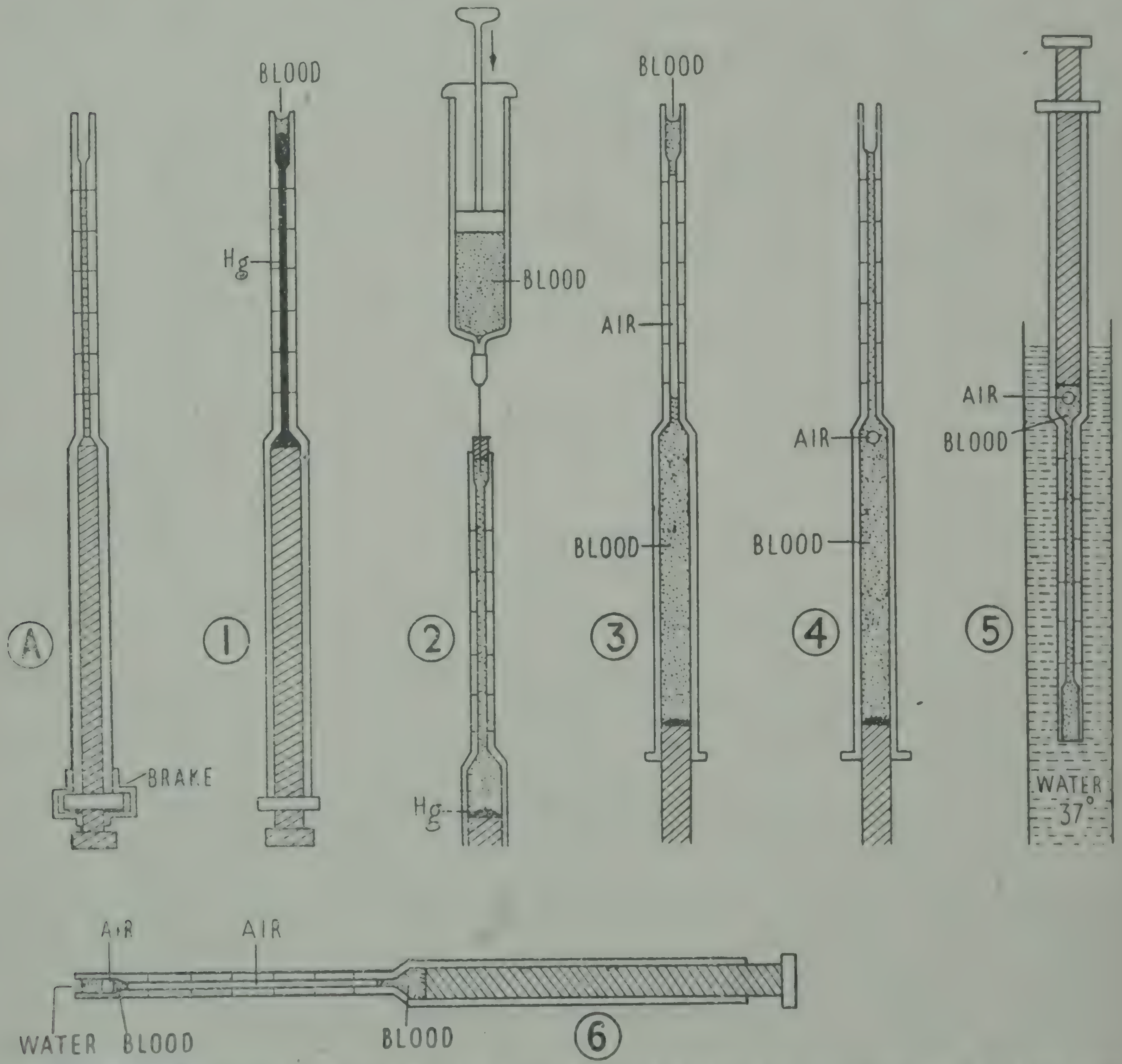


FIG. 499. Illustrating use of the Roughton-Scholander syringe for finding gas tensions in blood. A, the syringe : (1) Dead space filled with mercury : (2) Introducing the heparinised blood : (3) Bubble of alveolar air in capillary : (4) Bubble drawn into syringe : (5) After equilibration ; expelling the blood : (6) Bubble transferred to capillary for measurement and analysis. (After description by RILEY *et al.* *J. Biol. Chem.*, 1945, **161,** 621.)

in descending order: alveolar air, arterial blood, venous blood; and those of CO_2 in the reverse order.

The partial pressures of gases in the alveolar air are calculated from its analysis. The gaseous tensions in the blood, however, can only be got by finding the gaseous mixture with which the blood will be in equilibrium. If we know the percentage of any gas in this mixture, we know its partial pressure, and therefore the tension of that gas in the liquid. In all such experiments, the main difficulty is in ensuring that a sufficient volume of blood is equilibrated with a very small volume of the gas mixture. A relatively small volume of gas must be used in order not appreciably to alter the blood gas tensions. Krogh * therefore in his *microtonometer* adopted the ingenious device of limiting the volume of air to a small bubble, the superficial area of which is large in proportion to its bulk. This bubble, after equilibration in a stream of blood for some minutes, is transferred to a special capillary tube, in which its analysis can be carried out.

Modifications of this method enable the gas tensions in human arterial blood to be determined with reasonable accuracy.

The blood is drawn by arterial puncture and heparinised without air contact. One ml. of the blood is transferred to a specially modified syringe, bearing a graduated capillary having a terminal cup (Roughton-Scholander syringe)†,—again without air contact. A small bubble of alveolar air is drawn into the capillary and then into the syringe (Fig. 499). The syringe is then placed horizontally and rotated at 37° C. for five minutes to gain equilibrium: next the blood is nearly all expelled, the bubble driven into the capillary and measured. The bubble is then brought into contact with 4 per cent. NaOH, and again measured. Finally it is exposed to alkaline hydrosulphite to absorb oxygen and the residual nitrogen measured.

From this analysis of the gas bubble, its partial pressures, which are identical with the gas tensions in the blood, are calculated in the usual way.

The main aim in this method is to get the *oxygen* tension of the arterial blood. Carbon dioxide tensions are got indirectly, since it has been shown by simultaneous samplings of arterial blood and alveolar air (by Haldane and Priestley's method), that samples of alveolar air collected at the end of a normal *expiration* have approximately the same carbon dioxide pressure as that of the arterial blood, both during rest ‡ and during exercise.§

Rather varying figures have been given for the oxygen tension of arterial blood, but the more recent determinations tend to show that it is either equal to the partial pressure in the alveolar air,‖ or within a few mm. of it.¶ We may take 95 mm. as a probable figure at rest: this would correspond to about 98 per cent. saturation at 40 mm. CO_2, on the dissociation curve (Fig. 491). In exercise the O_2 tension in arterial blood may be as much as 16 mm. below that of alveolar air.‖ The value for CO_2 tension is generally agreed to be about 40 mm. at rest, or nearly the same as alveolar air.

There is also some doubt about the exact values of the pressures of both gases in the venous blood going to the lungs. An indirect method of finding them is by employing the lung itself as an aerotonometer. By a method already described (p. 569) a suitable gas mixture is held for about fifteen seconds in the lungs, and the alveolar air has then the gas tensions of pulmonary venous blood. By this means the tension of the oxygen in the venous blood was found to be about 40 mm. Hg, and that of the carbon dioxide 46 mm. Hg.

Another indirect method of finding the gas tensions in man is by reference to the oxygen or carbon dioxide dissociation curves. If we can find the oxygen and carbon dioxide contents of the bloods in question, reference to the dissociation curves will give their partial pressure.

---

* KROGH. *Skand. Arch. f. Physiol.*, 1908, **20**, 259.

† ROUGHTON and SCHOLANDER. *J. Biol. Chem.*, 1943, **148**, 541.

‡ BOCK and FIELD. *J. Biol. Chem.*, 1924, **62**, 269.

§ DILL, LAWRENCE, HURXTHAL and BOCK. *J. Biol. Chem.*, 1927, **74**, 313.

‖ COMROE and DRIPPS. *Amer. J. Physiol.*, 1944, **142**, 700; GALDSTON and WOLLACK. *Ibid.*, **1947**, **151**, 276.

¶ LILIENTHAL *et al.* *Amer. J. Physiol.*, 1946, **147**, 191.

The probable partial pressures and tensions concerned in respiratory interchange at rest are summarised in the following table :—

Approximate Partial Pressures in Inspired, Expired and Alveolar Air, and Tensions in Arterial and Venous Blood, at Barometric Pressures of 760 mm. Hg.

Gas	Pressure or Tension, mm. Hg.				
	Inspired Air	Expired Air (35·5° C.)	Alveolar Air (37° C.)	Arterial Blood	Venous Blood (Rest)
Oxygen	158·3	116	99	95	40
Carbon dioxide	0·3	30	39	40	46
Nitrogen	596·4	575	577	548	570
Water vapour	[*]5·0	39	45	47	47
Total pressure	760·0	760	760	730	703

It should be noted that the total pressure in blood is lower than that of the air.

As the venous blood enters the lungs, there is thus a difference of oxygen pressure of about 60 mm. Hg, which will tend to cause a flow of oxygen from alveolar air to blood, and a difference of 6 mm. Hg, tending to cause a flow of carbon dioxide from blood to alveolar air. Are these differences sufficient to account for the amounts of gases given off or taken up by the blood in its passage through the lungs ? The lungs have been estimated to contain 700×10^6. alveoli, each of a diameter of 0·2 mm., so that at medium distension the total surface amounts to at least 90 square metres (about 1,000 square feet). The lung capillaries have a total length of about 1,000 miles and an area of about 40 sq. m. when at rest.[†] Here some 60–95 c.c. of blood suddenly spreads out into a layer the thickness of which is not much more than that of one blood corpuscle, and is exposed to the air over this huge area, whence it is collected again into the pulmonary veins. The thickness of the tissue separating this layer of blood from the alveolar air is on the average 1·4 μ. If the "alveolar membrane" is reasonably permeable the chances for diffusion should be good, despite the fact that the blood only takes about a second to pass a lung capillary in resting states, and about one-third sec. in hard work.[†]

The diffusion constant (= c.c. O_2 passing from alveolar air to blood per mm. Hg tension difference of O_2, per min.) has been measured in man and found to vary from 21 at rest, to 62 in exercise.[‡] Let us call it 30 c.c./mm./min. Then it would be physically possible for $30 \times 60 = 1{,}800$ c.c. of oxygen to pass through into the blood per minute. The oxygen actually absorbed by a man at rest amounts to about 250 c.c. per minute, so that the physical conditions allow an ample margin for increase in the consumption of oxygen ; in fact, a difference of pressure of only 9 mm. would suffice to cause a passage of the 250 c.c. per minute which is required by the resting man. It is still easier to account for the passage of carbon dioxide in the reverse direction, although the pressure difference is only 6 mm. This gas diffuses through a wet membrane about twenty-five times as rapidly as oxygen, so that a difference of pressure between the venous blood and the alveolar air amounting to only 0·03 mm. Hg would suffice to cause a passage outwards of the 215 c.c. of CO_2 normally expired per minute.

It is evident that the limitation to the absorption of oxygen is really set by the capacity of the hæmoglobin to combine with it. From the oxygen dissociation curve of blood we see that the amount of oxygen which can be taken up by blood in the presence of the normal tension of carbon dioxide, *i.e.* 40 mm.

---

[*] Variable, according to temperature and relative humidity.

[†] Roughton. *Amer. J. Physiol.*, 1945, **143**, 621.

[‡] M. Krogh. *J. Physiol.*, 1915, **49**, 271 ; Lilienthal *et al.* *Amer. J. Physiol.*, 1946, **147**, 199.

Hg, begins to diminish very rapidly when the pressure of the oxygen falls below 50 mm. Hg. Thus, at 40 mm. oxygen pressure, and a carbon dioxide tension of 40 mm., blood is about 75 per cent. saturated, and at 30 mm. it is only 60 per cent. saturated. Under normal circumstances, the blood leaves the lungs about 98 per cent. saturated with oxygen. If the saturation fell to 60 per cent. we should expect to obtain evidence of failure of oxygen supply to the tissues. Probably the oxygen tension in the alveoli can sink to between 30 and 35 mm. Hg before any signs of oxygen lack make their appearance. These results were obtained by exposing a man, in a state of complete rest, to reduced pressure in an air-chamber. Under these conditions, the slightest muscular exertion would cause distress from deficient oxygen supply. The exact percentage of oxygen in the inspired air, which would give an alveolar oxygen tension of 30 to 35 mm., varies with the depth of respiration. Thus, with shallow respiratory movements the pressure may sink to 35 mm. Hg when the inspired air contains as much as 12 per cent. oxygen. If the movements be deeper, the oxygen content of inspired air may be reduced to 9 or 10 per cent. before respiratory distress is observed.

The view that, in the interchange of gases in the lungs, the membrane between the blood and the alveolar air plays simply a passive part is now almost universally accepted since experiments show conclusively that the difference between the tensions in the alveoli, and in the blood, respectively, is always such as to allow of the passage by diffusion of oxygen inwards and carbon dioxide outwards from the blood. Moreover, the structure of the pulmonary epithelium lends no support to the view that it acts as a secreting membrane. In mammals the cells are of two kinds, viz. small granular nucleated cells lying in the interstices of the capillaries, and larger, extremely thin, structureless plates, *without nuclei*, covering the capillaries.

The view that the pulmonary epithelium actively secretes oxygen was held by Bohr,* and Haldane † held that, at all events under circumstances when the oxygen pressure is low in the alveoli, oxygen may be secreted, so as to attain a *higher* tension in the arterial blood than in the alveolar air. Thus, at high altitudes, Haldane claimed that such a secretion could be demonstrated if the subject is given air containing a low percentage of carbon monoxide to breathe. Under these circumstances the blood is found to contain less CO than when shaken with the same alveolar air (containing CO) *in vitro* ; *i.e.* the uptake of oxygen had been favoured, and the uptake of CO therefore impeded, in the lungs.

A reinvestigation by Barcroft,‡ of the tensions of oxygen in the blood under such conditions, and using direct methods, failed to confirm Haldane's conclusions.

ABSORPTION OF GAS FROM THE PLEURAL CAVITY. If air is introduced into the pleural space in sufficient amount to cause an artificial pneumothorax (*e.g.* in treatment of pulmonary tuberculosis) the lung on that side collapses. The air so introduced is soon altered in composition and comes into approximate equilibrium with the blood in the lung capillaries by diffusion through the visceral pleura. This result having been achieved, it might be expected that the gas would stay there indefinitely, but actually it disappears by absorption into the venous blood entering the lung capillaries. The reason for this is that, whereas the total pressure of the air in the pleural space is practically atmospheric, that in the venous blood (*v.* table, p. 740) is some 57 mm. Hg below atmospheric. Hence all the gases finally diffuse into the blood.

EXCHANGE OF GASES IN THE TISSUES

The considerations which determine the exchange of gases between air and blood in the lungs are also operative in determining the gaseous exchanges which take place during the circulation of the blood through the capillaries of the tissues. Blood enters the capillary network as arterial blood, and having, therefore, an oxygen pressure of the order of 95 mm. Hg and a carbon dioxide pressure of the order of 40 mm. Hg. In the tissues the oxygen pressure is lower and the CO_2 pressure higher than in the arterial blood.

The rate of passage of oxygen out of the capillaries into the tissues must be proportional to the difference of oxygen tension between the capillary blood and the tissue cells. If Pb is the oxygen pressure in the capillary blood and Pt that in the tissues, the rate of flow of oxygen from capillaries to tissues must be proportional to Pb—Pt. If the blood is to lose oxygen during the whole of this passage through a tissue, we may take for Pb the tension of the oxygen in the venous blood as it leaves the tissue. This is, say, 30 mm. Hg. Pt, the tension of oxygen in the tissues, may be determined in various ways,* but is usually considerably lower than Pb.

---

* BOHR. *Skand. Arch. f. Physiol.*, 1890, **2**, 236.

† DOUGLAS and HALDANE. *J. Physiol.*, 1912, **44**, 305.

‡ BARCROFT *et al.* *J. Physiol.*, 1920, **53**, 450.

Krogh † determined the rate of diffusion of oxygen through different thicknesses of various tissues (connective tissue, muscle, etc.) under varying pressures. Then, partly on the living muscle, partly on injected specimens, he estimated the relation of the number and size of the capillaries to the intercapillary muscular tissues. He found that extremely small pressure differences were necessary to supply the muscle fibres with oxygen. Some of his results on the guinea-pig are shown in the accompanying Table.

Guinea-pig muscle	c.c.O_2 per minute per 100 c.c. tissue	Capillaries per mm.^{2}	Pb − Pt mm. Hg.	Total surface capillary in 1 c.c. muscle	Capacity of capillaries in 100 c.c. muscle
Rest	0·5	31	45	3 cm.^{2}	0·02 c.c.
	0·5	95	12	8 ,,	0·06 ,,
	0·5	270	3	32 ,,	0·3 ,,
Massage	0·5	1,400	0·04	200 ,,	2·8 ,,
Work	5	2,500	1·4	390 ,,	5·5 ,,
Maximum circulation	10	3,000	1·2	750 ,,	15 ,,

The first thing that strikes us in this Table is the enormous difference between the capillary circulation of resting and that of active muscle. In the resting muscle the majority of the capillaries are empty and collapsed, so that relatively large volumes of muscle intervene between the few capillaries in which the circulation of blood is proceeding. Under these conditions, the oxygen pressure head necessary to supply the total oxygen needed by the whole of the muscle, *e.g.* 45 mm. Hg, may be greater than the venous oxygen tension, so that in parts of the muscle the oxygen tension may be zero: capillaries in these anoxic parts then open up, and others close. After massage, a number of new capillaries open; the number is still further increased by work, so that there may be a hundredfold increase in the number of capillaries in every square millimetre of a cross-section of the muscle. Under these conditions, the passage of oxygen from the capillaries is so facilitated that the oxygen pressure in the muscle tissues becomes practically equal to that of the blood. It would appear that, so far as the supply of oxygen to the muscle is concerned, the increase in the capillary area during muscular exercise is far ahead of the actual needs of the muscle.

Under normal circumstances the blood never stays in the capillaries long enough to lose the whole of its oxygen. The fraction of the arterial oxygen abstracted, which is called the *coefficient of oxygen utilisation*, varies from tissue to tissue, and from time to time. In states of activity, the venous blood may be more than usually deoxygenated, but generally not more than 0·4 of the arterial oxygen is utilised. If, however, the further supply of oxygen to the blood be prevented, as in asphyxia, the last traces of oxygen may disappear from the blood.

As a result of the oxidative changes in the tissues, carbon dioxide is produced, and the tension of this gas in the tissues therefore rises.

The alteration in reaction of the blood, caused by a rise in CO_2 tension or by the presence of small amounts of lactic acid, markedly quickens the rate at which oxyhæmoglobin gives up its oxygen, as is shown in Fig. 500.

As Barcroft has pointed out,‡ in cold-blooded animals the hæmoglobin is different in its properties from that of warm-blooded animals, being less affected by temperature changes. In them, too, the dissociation of oxyhæmoglobin must be largely conditioned by the rise of carbon dioxide tension in the tissues, since at the normal temperature of these animals the evolution of oxygen from hæmoglobin would be slow.

---

* VERZÁR. *J. Physiol.*, 1912, **45**, 39.

† KROGH. "The Anatomy and Physiology of the Capillaries." (Yale Press, 1929.)

‡ BARCROFT. *Physiol. Rev.*, 1925, **5**, 596.

The determination of the pressures of oxygen and carbon dioxide in the tissues is a matter of some difficulty. They may be approximately determined by taking the tensions of the gases in secretions, such as the urine or bile or, more generally, by injecting air beneath the skin, or into body cavities, etc., and analysing it after the lapse of a suitable interval to allow of equilibration.* It is found to be from 45 to 55 mm. Hg for CO_2 (14 to 23 mm. for cold-blooded animals),† and since in venous blood it is rarely above 50 mm. Hg, there is a gradient of partial pressure of CO_2 from tissues to blood, *viâ* the tissue fluid. The gas therefore passes in that direction by simple diffusion, and the oxygen in like manner passes in the opposite direction, also by diffusion. We have already seen that the oxygen pressure may vary from one tissue to another, or in the same tissue, from time to time. It is certain, however, that the oxygen pressure in the tissues is always lower than that in the arterial blood and tends to vary inversely as the activity of the tissue. By the method mentioned above the oxygen pressure of the body cavities and of subcutaneous tissue appears to range from 20 to 45 mm. Hg. In muscle it may be lower than this, especially during activity. Under the skin, on the contrary, O_2-tension is raised during exercise, because the cutaneous vessels are dilated.

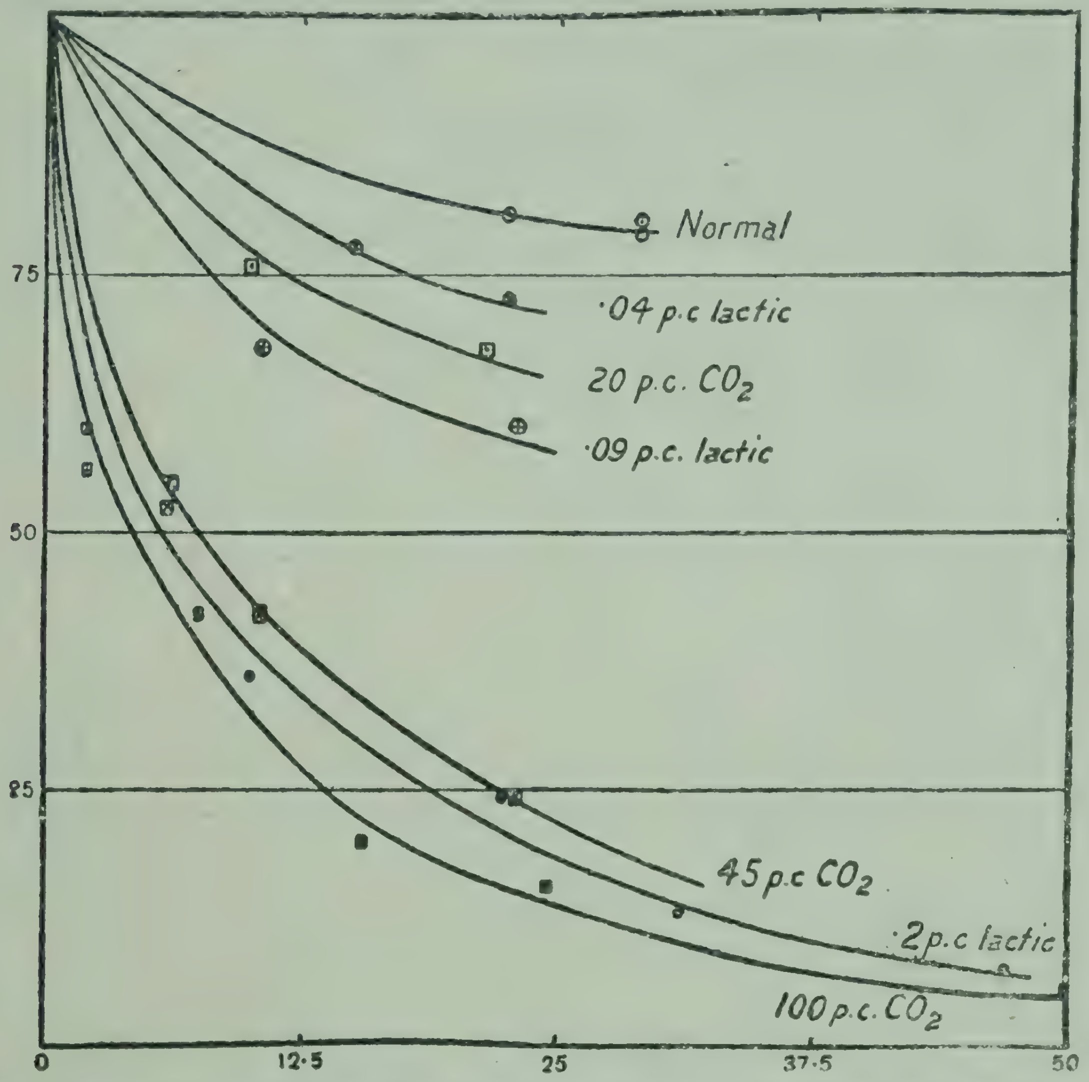


FIG. 500. Curves showing the Rate at which Arterial Blood is Reduced on Bubbling through it a gas free from Oxygen, and the effect on the rate of the presence of CO_2 and of Lactic Acid.

Ordinates = percentage saturation of oxyhæmoglobin. Abscissæ = time in minutes. (MATHISON. *J. Physiol.*, 1911, **43**, 347.)

Increasing or reducing the oxygen tension in the air breathed causes a change in the same direction in both the oxygen and the CO_2-tensions in the tissues. The CO_2-pressure is increased when the oxygen-pressure is increased because CO_2 is then less readily taken up by blood and so accumulates in the tissues.‡

---

* CAMPBELL. *J. Physiol.*, 1924, **59**, 1.
† CAMPBELL. *J. Physiol.*, 1926, **61**, 248.
‡ CAMPBELL. *J. Physiol.*, 1925, **60**, 20.

In muscles that exhibit strong and repeated contractions there is a modified form of hæmoglobin present in the cytoplasm of the fibres. It is called *myohæmoglobin*, or *myoglobin*, and serves a useful purpose in facilitating the transport and storage of oxygen. It differs from blood hæmoglobin in its absorption spectrum, in having only one atom of iron per molecule, in having a relatively small affinity for CO, but especially in showing a very small Bohr effect (p. 729), and in having an oxygen dissociation curve that is a rectangular hyperbola. Consequently, it will take up oxygen much more readily at low pressures than will blood hæmoglobin, so that it will easily deprive the latter of oxygen, particularly if CO_2 is also present. It also has a higher velocity of combination with oxygen than has blood hæmoglobin, so that it can act as a quickly filled store of oxygen for the muscle fibres. It has been calculated that the myoglobin of heart muscle can hold enough oxygen for about 7 secs. of moderate activity, or about half a second of maximal activity of the heart muscle, and so should be able, with its quick rate of oxygen uptake, to keep the tissue always supplied with oxygen. It will be seen from the dissociation curves of Fig. 501 that even at the venous oxygen tension of 40 mm. the myoglobin is still 94 per cent. saturated, whereas the blood Hb is only 66 per cent. saturated. Hence the tissue could suck oxygen out even from the venous blood.

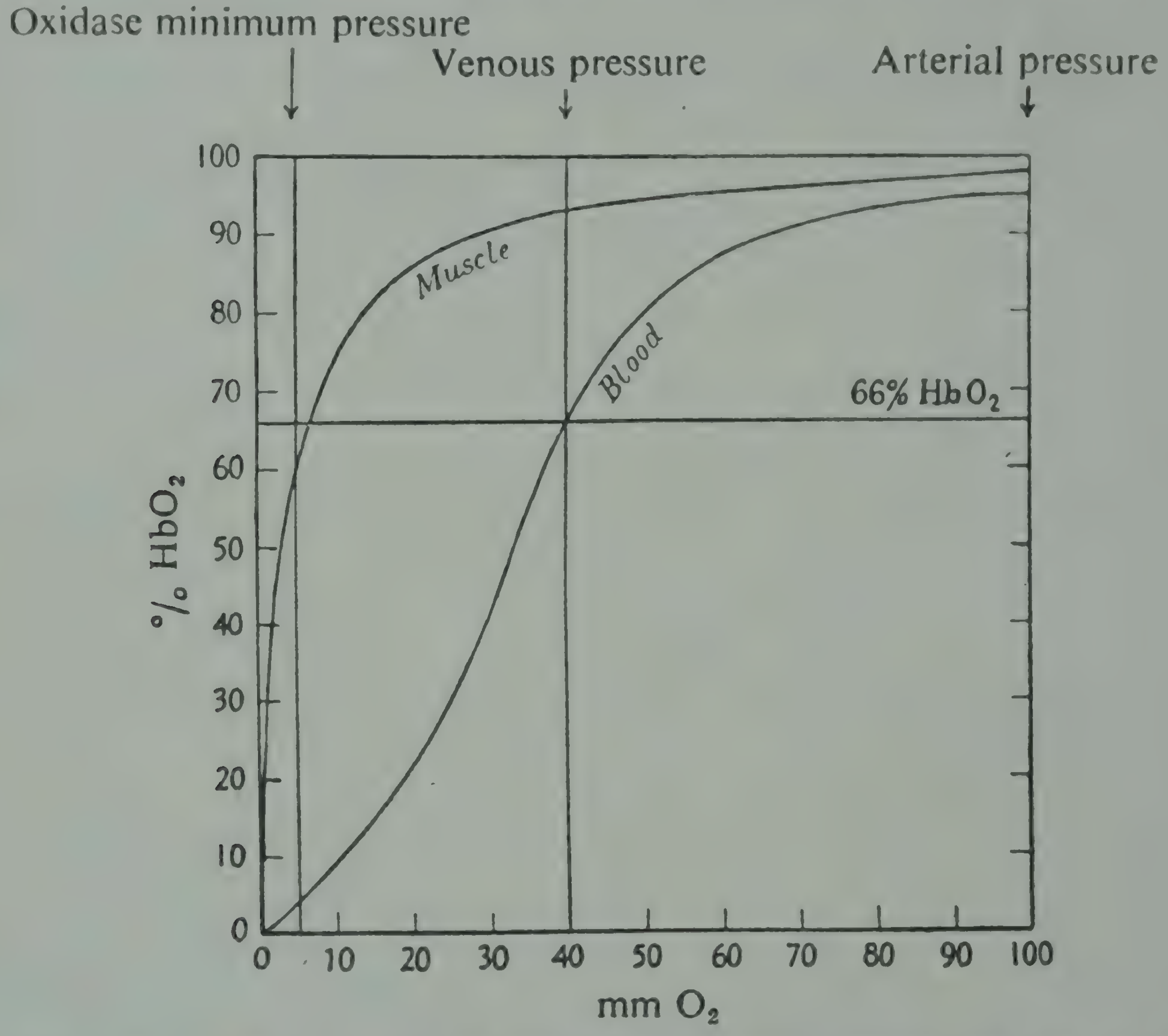


FIG. 501. Dissociation curves of muscle hæmoglobin (myoglobin) and of blood, showing relation to minimum oxygen pressure required by the tissue oxidases. (After R. HILL. *Proc. R. S.*, 1936, **120B**, 472.)

As a store of oxygen, however, the myoglobin would be of no use if it held its oxygen so firmly that the active systems of the tissue cells were unable to withdraw it. It will be seen from Fig. 501 that this is not the case, for the tissue oxidases can function at oxygen pressures of 5 mm. or less, so that much of the myoglobin can be readily reduced.

Use has been made by Millikan * of the presence of myoglobin in the muscle fibres for studying the utilisation of oxygen by living muscles. Light is passed through the living soleus muscle of the cat, and falls on to a differential photocell colorimeter, by means of which a continuous record of the degree of oxygenation of the myoglobin is made on photographic paper. When the muscle was tetanised the myoglobin showed reduction in 0·2 sec., and had lost up to 40 per cent. of its oxygen in the first second, even when the blood supply was intact. On relaxation the oxygen was restored. These facts are of importance in emphasising the speed with which reactions occur in the mammalian tissues with intact circulation.

---

* MILLIKAN. *Proc. R. S.*, 1937, **123B**, 218.

Tissue Respiration. The enormous avidity of the tissues for oxygen is shown by the following experiment (Ehrlich).[*] If a saturated solution of methylene blue be injected into the circulation of a living animal and the animal be killed ten minutes later, it is found, on first opening the body, that most of the organs present their natural colour, although the blood is of a dark blue colour. On exposure to the atmosphere, all the organs acquire a vivid blue colour. This phenomenon is due to an intense reducing activity of the tissues, whereby hydrogen is added to the methylene blue with the production of a colourless reduction product, which on re-exposure to oxygen is oxidised, by loss of hydrogen, into the blue compound. It is shown by surviving minced tissues as well as *in vivo*. In this reaction the tissue obtains oxygen, for carrying out its oxidative processes, by the decomposition of water, the oxygen from the water being transferred to an oxidisable substance (oxygen acceptor), provided there be present a substance, such as the methylene blue (hydrogen acceptor), which can take up the active hydrogen (*v.* p. 98).

All tissues use oxygen and form CO_2, in amounts varying with their state of activity, since their energy is derived from combustion processes. The heart, brain and liver are among the most actively respiring organs, but during muscular work the contracting muscles also consume large amounts of oxygen. The intake of oxygen and output of CO_2 by the lungs represents the sum total of the gaseous exchanges of all the individual tissues of the body. The lungs themselves have a small oxygen usage and CO_2 production, but, apart from this, all the oxygen taken in from the air is carried away in the arterial blood, and all the CO_2 expired is carried to the lungs by the venous blood. In other words, the lungs do not effect, or finish off, combustions on behalf of the other tissues.[†]

The respiratory exchanges can be studied (i) in the whole body, by measuring the exchanges effected in the air in the lungs. This aspect of the subject will be dealt with in Chapter XLI, and need not be further discussed here.

(ii) in a whole organ or mass of tissue. This can be done by determining the rate of blood flow through it, and by measuring the arterio-venous oxygen and/or CO_2 differences.

(iii) in thin slices of tissue, or membranes, such as the diaphragm, retina, etc. Since a circulation of blood is not indispensable for carrying out of tissue respiration, provided that oxygen and CO_2 exchanges are provided for in some other way, advantage is taken of the fact that thin slices of fresh surviving tissue still respire when placed in oxygen-rich atmospheres. The study of the respiration of such tissue slices by means of the various types of microrespirometer (*e.g.*, that mentioned on p. 146) is now much in vogue. The slices must be thin enough to ensure an adequate oxygen supply, and must be suspended in a suitable isotonic fluid, to which various substances of metabolic importance can be added, and the oxygen uptake and CO_2 production followed. To offset uptake of water by the slices, it is common to express the results as so-called Q_{O_2} values, *i.e.* μl (= c.mm.) O_2 consumed per mg. dry weight of tissue per hour. To facilitate comparison of results it is also a practice in microrespirometry to express all quantities of substances, solids, liquids or gases, as though they were in gaseous form, and so to express 1 millimol. as = 22,400 μl.

(iv) in finely ground tissue pastes, in which most of the cells are destroyed. As a rule it is found that the greater the cell destruction the smaller is the survival respiration : one explanation of this is that cellular structure implies the existence of intracellular apparatus in which the chemical changes are brought about. If we minced up an automobile, we should hardly expect it to display any trace of its normal functions, particularly if all the gasoline and oil had run away. Another explanation of the reduced respiration of ground-up tissue is that substances essential for the combustion process, either as fuel or as catalysts, have diffused away from the cells. In support of this, it has been shown that when certain substances, believed to be metabolites, are added to the paste, a normal rate of respiration is resumed. Does this, in fact, mean that the normal function and normal respiratory process is reinstated, or is it more comparable with pouring gasoline or oil over the red-hot debris of our automobile ? Perhaps a little of both !

---

[*] Ehrlich. "Das Sauerstoff-Bedürfniss des Organismus" (Berlin: Hirschwald, 1885); *Dtsch. med. Woch.*, 1886, No. 4. For a discussion regarding the physico-chemical significance of the reduction of dyestuffs, *v.* Mansfeld Clark *et al.*: *Hyg. Lab. Bull. U.S. Publ. Health Service*, 1928, 151; *Chem. Reviews*, 1925, **2**, 127.

[†] Evans and Starling. *J. Physiol.*, 1913, **46**, 413.

(v) in isolated enzyme systems extracted from the tissue, and acting on appropriate substrates, as already explained on pp. 98 and 99).

It will be evident that these various methods of approach give us, in the order named, an increasing power of analysis of the biochemical processes concerned in tissue respiration, but at the same time conditions which depart more and more from those of the normally functioning cells of the tissue in question. All available methods must, therefore, be employed, and the evidence which they provide carefully sifted.

As an example of the application of the various methods we may take first the gaseous metabolism of the heart. We can determine this in the actively working organ by the use of the heart-lung preparation. First, we can determine the oxygen uptake and CO_2 output of the HLP by way of the lungs, and then subtract the quantities due to the metabolism of the lungs themselves. In this way we find that the heart of the dog, while doing moderate amounts of work, uses about 350 c.c. O_2/100 g./hr. Next, we can collect and measure the blood flowing through the coronary circulation during a given time, find the oxygen content of it, and of the arterial blood, and so calculate how much oxygen the heart used. For instance, if the heart weighed 67 g., the total venous blood flow from the heart was 42·5 c.c./min., and the arterio-venous oxygen difference 0·061 c.c. O_2/c.c. blood, then the oxygen usage of the heart would be :

$$\frac{42{\cdot}5 \times 0{\cdot}061 \times 60 \times 100}{67} = 232 \text{ cc./100 g./hr.}$$

By the study of a tissue pulp in oxygen in a microrespirometer, heart tissue weighing 375 mg. in one hour consumed 711 μl. O_2, which = 189 c.c./100 g./hr. On adding to the pulp a 0·02 M. solution of pyruvate, the consumption rose to 1134 μl./hr., *i.e.* to 302 c.c./100 g./hr.* This observation confirmed an earlier one, made with the heart-lung preparation, in which the heart was doing a steady rate of work. The heart was at first using oxygen at the rate of 360 c.c./100 g./hr. : addition to the circulating blood of 0·8 g. of sodium pyruvate raised the oxygen usage to 435 c.c./100 g./hr.† In both experiments disappearance of pyruvate was proved by analysis. We may infer that, especially in the case of the tissue pulp, the pyruvate provided combustible material which was lacking, and so brought up the oxygen usage of the pulp to much the same figure as that shown by the intact heart tissue doing mechanical work.

As another instance of the use of the microrespirometer in pointing to possible combustible substrates, we may take the following figures, which show the combustion of glucose and of pyruvic acid by rat brain cortex in a phosphate buffer solution ‡ :—

Addition to medium	Q_{O_2}
None.	2.98 ± 0.18
Glucose, 0.01 M.	12.20 ± 0.33
Pyruvate, 0.02 M.	11.10 ± 0.65

---

* Smyth. *Bioch. J.*, 1940, **34**, 1046.

† Braun-Medendez *et al.* *Q. J. Exp. Physiol.*, 1939, **29**, 91.

‡ Jowett and Quastel. *Bioch. J.*, 1937, **31**, 565.

It is becoming increasingly clear that the biochemical processes involved in tissue respiration are inseparably connected with those of metabolism in general, and that, between the entry of oxygen into a living cell and the final output of CO_2 from it there lie concealed many happenings which appear to us, in our limited way of thinking, to be inordinately complex and mysterious.

We have already seen in Chapters IX and X something of the biochemical mechanisms of tissue oxidations, and learned that this is a result of a complicated chain of reactions, which probably differ in detail from tissue to tissue, though on the whole conforming to a common plan. There are also in all probability, alternative paths by which oxidation may take place, in a given tissue, according to the conditions.

By way of further illustration and recapitulation of data given in Chapters IX and X, we may take the following examples of experiments illustrating processes that occur in heart muscle.

(i) By the use of the microrespirometer, thin slices of heart suspended in a buffer solution at *p*H 7·3 and 37° C. are found to use oxygen at a regular rate. If a trace of HCN is added to the solution, the oxygen uptake ceases.

(ii) Fresh heart tissue is minced, washed with water, ground thoroughly with M/25 phosphate buffer, *p*H 7, and centrifuged. The cloudy fluid is treated with M/5 acetate buffer, *p*H 4·5, centrifuged again, and the solid suspended in M/10 phosphate buffer, *p*H 7·3. This suspension contains powerful oxidases. If we place, say, 0·5 c.c. of this in a microrespirometer, with 2·3 c.c. of M/10 phosphate buffer, *p*H 7·3, and 0·3 c.c. of 6 per cent. *p*-phenylene diamine, and shake, a steady oxygen usage is seen, say, 50 μl. in ten minutes. The *p*-phenylene diamine is being oxidised, and oxygen is absorbed.

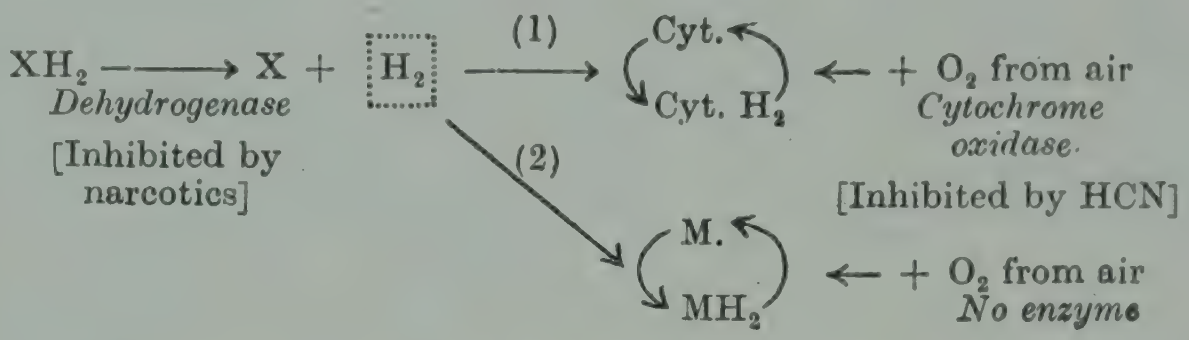
Now let us repeat the above experiment, but with the addition to the contents of the microrespirometer of a few mg. of cytochrome (prepared from heart muscle): it is now seen that the oxidation of the *p*-phenylene diamine proceeds much more rapidly than before, say about 120 μl. in ten minutes. If a trace of cyanide be now added the oxygen usage stops altogether. The addition of cytochrome was effective because in the preparation of the oxidase much of the original cytochrome of the tissue was lost.

What happens to the substrate (in this case the *p*-phenylene diamine, but in the living tissue some other substance, which is to be oxidised) is that it hands hydrogen to cytochrome under the action of a dehydrogenase, and so is oxidised, while the cytochrome is reduced. The reduced cytochrome is then re-oxidised, with absorption of oxygen, by another enzyme, cytochrome oxidase. This oxidase is inhibited by cyanide, so that the oxygen usage ceases. Presumably a similar thing happens in the tissue slice when cyanide is added.

(iii) If we mix in an evacuated tube 0·3 c.c. of M/2 sodium succinate, 1 c.c. of a washed heart suspension, 0·5 c.c. of M/1,000 methylene blue, and 1·5 c.c. of M/10 buffer, *p*H 7·3, and keep at 37° C., the blue colour will be discharged in a few minutes, but returns at once if air is admitted. If the succinate is omitted, the methylene blue is not reduced by the washed heart tissue *in vacuo*. Here the substrate is succinic acid, the enzyme a dehydrogenase which oxidises the succinate to fumarate by removal of hydrogen: the hydrogen is taken up by the methylene blue, which is decolorised, but on admitting air is reoxidised by the free oxygen. If we perform the experiment in a microrespirometer, we find that succinate plus heart tissue uses oxygen, and that the addition of cyanide stops this usage, but if methylene blue is also added, the oxygen usage goes on in spite of the presence of the cyanide. The hydrogen acceptor now is not cytochrome, but methylene blue (and in the body might well be some cell constituent): cyanide does not inhibit either the action of the dehydrogenase, nor the autoxidation of the methylene blue, and so, if methylene blue is present, the oxygen usage can continue in spite of the presence of the cyanide.

(iv) Although in the above example the cytochrome oxidase plays no part in the oxidation of succinate, it can do so under the proper conditions. If a mixture of oxidase and cytochrome is examined spectroscopically, the indefinite absorption of oxidised cytochrome is all that is seen. If now a few drops of succinate solution are added, the distinct absorption bands of reduced cytochrome immediately appear (Fig. 332, p. 514) : on shaking with air, they vanish again, to reappear after a few seconds further standing. If, before the succinate, a few drops of a solution of urethane be added, no reduction is produced by succinate : if, on the other hand, succinate and oxidase-cytochrome mixture be made, and cyanide added, the reduced cytochrome remains unchanged when the solution is shaken with air.

These various facts can be brought together in the following scheme, in which XH_2 is the substrate to be oxidised (*e.g.*, succinate), X is the oxidised substrate, M is methylene blue, MH_2 the reduced "methylene white," and Cyt. is oxidised cytochrome.



We have already seen (pp. 98–99) that the dehydrogenases may be quite complicated systems, consisting of an enzymic protein part, and one or more co-enzymes, which may themselves at certain stages act as substrates for the enzyme. For example, in the reversible oxidation of lactic to pyruvic acid, we may have a complex system consisting of a dehydrogenase enzyme, DPN, (*v.* p. 93), flavo-protein (p. 94), and a carrier, such as methylene blue. First, the lactate is oxidised and the DPN reduced : then the reduced $DPNH_2$ is reoxidised to DPN, and the leucoflavin is reduced to the leuco form : then the leuco flavoprotein is oxidised again, and the methylene blue reduced to its leuco form : finally this takes up oxygen and is converted into the blue again (*v.* p. 99). Here it is not easy to draw the line between substrate and catalyst, and the same is true of many types of biochemical change, when we go into detail.

So far, we have spoken mainly of the uptake of oxygen, and have had little to say about the processes by which carbon dioxide is produced. Some hint of the sort of chemical change which may be responsible for this is given on p. 142, but there are other possibilities which are too complex to be considered here.

For Reference

Campbell (1931). Gas Tensions in the Tissues. *Physiol. Rev.*, **11,** 1.
Dixon (1934). Manometric Methods. (London : Cambridge Univ. Press.)
Elliott (1941). Intermediary Metabolites and Respiratory Catalysis. *Physiol. Rev.*, **21,** 267.
Meldrum (1934). Cellular Respiration. (London : Methuen.)
Meyerhof *et al* (1942). A Symposium on Respiratory Enzymes. (Univ. Wisconsin Press.)
Tobias (1943). Microrespiration Techniques. *Physiol. Rev.*, **23,** 51.

CHAPTER XL

THE REGULATION OF THE RESPIRATORY MOVEMENTS

Each movement of inspiration involves the co-ordinated activity of a number of muscles, and the extent to which they contract will determine the depth of the inspiration. Similarly, if expiration is to take place, they must simultaneously cease to act. The rhythm and extent of the alternate contractions and relaxations of the respiratory muscles are so regulated that the total ventilation is sufficient to keep the composition of the gas in the alveoli at a practically constant level.

Each act of inspiration involves a discharge along a number of nerves, *e.g.* the facial to the muscles moving the alæ nasi, the vagus to the muscles of the larynx, the branches of the cervical and brachial nerves to the muscles of the neck, the phrenic nerves to the diaphragm, and the thoracic nerves to the intercostal muscles. These nerves are derived from motor nerve cells, situated at various levels in the medulla and spinal cord. In each act of inspiration or expiration the activities of all these groups of cells must be correlated.

RESPIRATORY CENTRES

The co-ordination is effected by the subjection of these motor nuclei to the control of specialised neurone pools in the central nervous system, which act as receiving centres for afferent impressions from the lungs and various other parts of the body, and which are also sensitive to changes in the composition of the blood circulating through their vessels. These dominant parts are called *respiratory centres.* If the spinal cord be cut across at the seventh cervical nerve roots, the respiratory action of the intercostal and abdominal muscles ceases permanently, although respiration is still continued by the diaphragm and the other muscles supplied by nerves leaving the central nervous system above the point of section. It has been known since Galen that division of the cord at the first or second cervical nerves abolishes the action of the diaphragm as well, though ineffectual movements of the muscles supplied by the facial, vagus, and spinal accessory nerves continue. A section of the brain stem through the mid-brain leaves the respiratory movements unaltered. We must conclude from these experiments that the respiratory motor neurones of the cord are subject to control by impulses originating in the hind brain, and transmitted therefrom down the spinal cord.

Many experiments have been made with the idea of locating the position of the medullary respiratory centre more accurately. Legallois and Flourens * described it as near the apex of the *calamus scriptorius.* The centre is bilateral, each centre presiding more especially over the muscles of the same side of the body, so that longitudinal section at the middle line does not destroy the respiratory movements † ; it lies in the reticular formation, in close connection with the central terminations of the vagus nerves.

Accurate delimitation of the respiratory centre has been made in the cat.‡ It is probably similar in position in man. By electrical stimulation of known regions of the brain stem, respiratory responses were obtained from two neighbouring neurone pools, one giving inspiration and one expiration, so showing clearly for the first time that there is both an *inspiratory and an expiratory centre.*

---

* Historical account *v.* Langendorff, in Nagel's *Hdbch. d. Physiol. d. Menschen,* 1909, IV, 334.

† Porter. *J. Physiol.,* 1895, **17**, 455.

‡ Pitts, Magoun and Ranson. *Amer. J. Physiol.,* 1939, **126**, 673.

The *inspiratory centre* is in the ventral reticular formation immediately over the cephalic four-fifths of the inferior olive (Fig. 502). When this region is stimulated by rapid faradic shocks, regular breathing is suspended and a deep inspiration, involving both thorax and diaphragm, occurs, and the breath is held in that position as long as the stimulus is kept up (Fig. 503, A).

The *expiratory centre* lies slightly cephalic and dorsal to the inspiratory centre. When it is stimulated a full expiratory standstill occurs (relaxation of diaphragm and other inspiratory muscles, and often also a contraction of expiratory muscles) and this expiratory position is kept up (within limits) for as long as the stimulation is maintained (Fig. 503, B). On prolonged stimulation (over about 3 mins.), however, inspiration breaks through.

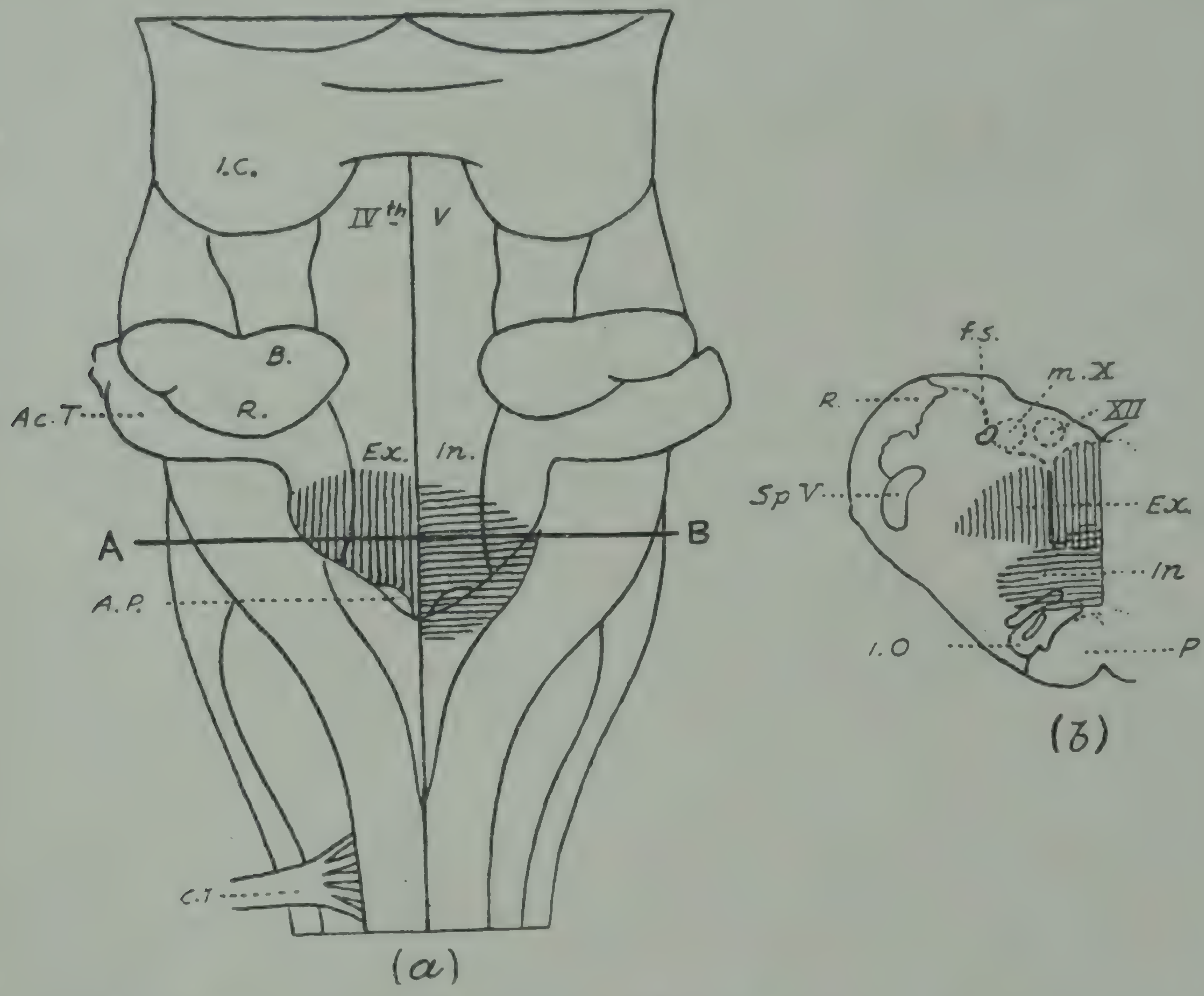


FIG. 502. Diagram showing the positions of the medullary respiratory centres in the cat.

(*a*) Dorsal view of brain stem after removal of cerebellum. Shaded areas show projection of respiratory centres on to floor of IVth ventricle. Vertical shading Ex. on left, area which on stimulation causes maximal expiration; horizontal shading In, on right, area which on stimulation causes maximal inspiration.

(*b*) Section along line A.B. of (*a*) showing position of areas projected, and showing that the expiratory centre lies dorsally to the inspiratory centre.

Ac.T., acoustic tubercle; A.P., area postrema; B., brachium conjunctivum; C.1, first cervical root; R., restiform body; I.O., inferior olive; Sp.V., spinal root of Vth cranial nerve; m.X., motor root of vagus; XII., hypoglossal nucleus; P., pyramid; Ex., In., expiratory and inspiratory centres; f.s., fasciculus solitarius.

(Modified from PITTS, MAGOUN and RANSON. *Amer. J. Physiol.*, 1939, **126**, 673.)

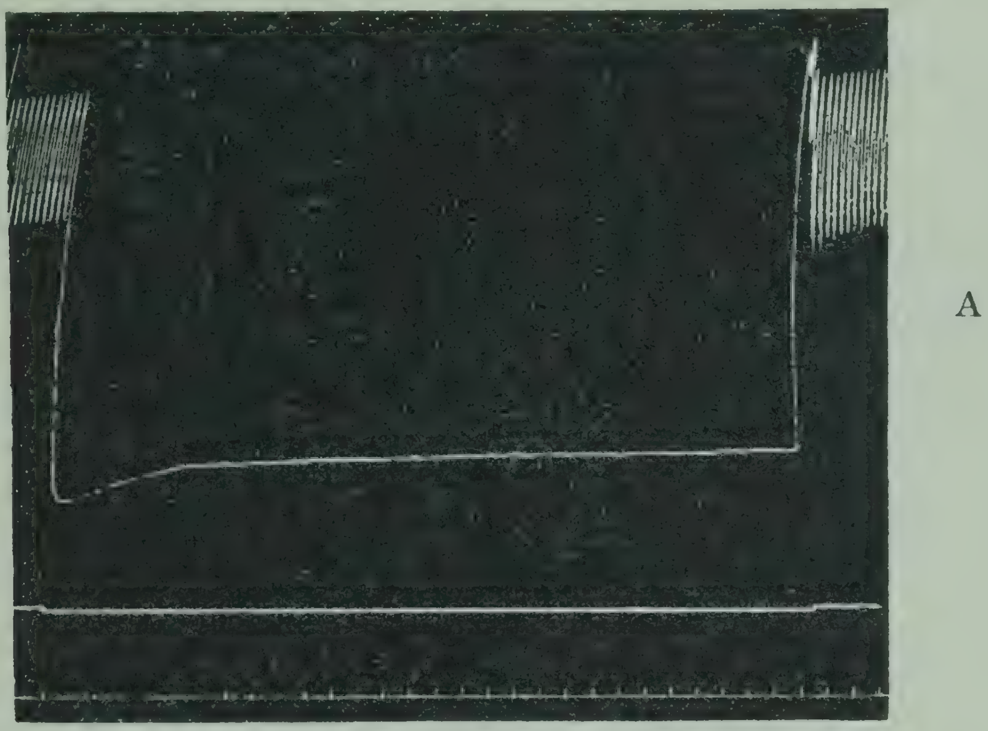
The inspiratory centre is the dominant one if both are equally stimulated together: by alternate stimulation of the two centres regular respirations of any desired depth or frequency can be obtained.

The neurones of the two centres are freely connected (Fig. 506), so that irradiation over the whole of a centre follows if only a small part of it is excited;

moreover, the other centre of opposite function is also affected, but in the opposite direction. Thus a stimulus applied to the inspiratory centre spreads over the whole of that centre, causing excitation, and also spreads to the expiratory centre, causing it to be inhibited. This reciprocal action between centres recalls the "half-centre" functions of limb locomotor centres advanced by Graham Brown (p. 266).

It is claimed by some physiologists that the primary neurones that are concerned with inspiration and expiration are not segregated, but are intermixed within a common centre, and that, owing to the reciprocal action mentioned above and to sundry incidental factors, the clear-cut results of inspiration or expiration on excitation necessarily follow. They claim that the electrical stimulus may be affecting afferent or efferent connections, rather than the neurocytons of the respiratory centre.*

Like all complex co-ordinated movements, the act of respiration, though perhaps capable of being carried out under the control of neurone pools, such as the medullary centres in question, is subject to various refinements of adjustment which can be brought into play through the intervention of superior centres.† These centres have often been grouped with the medullary centre as representing specialised departments of a central respiratory control mechanism, as by Lumsden, for example.‡



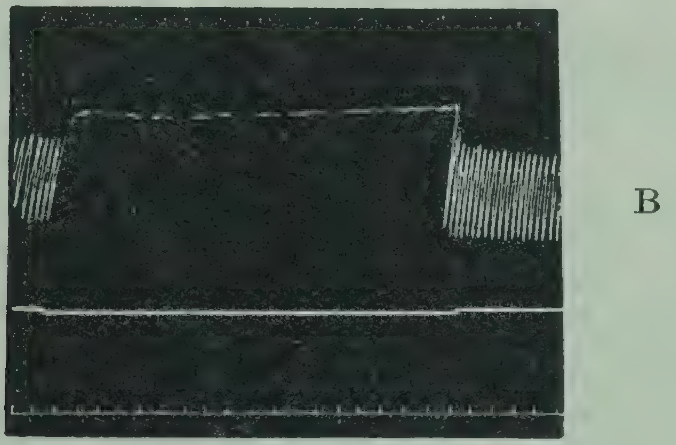


FIG. 503. Spirometer tracings of respiratory reactions to stimulation of respiratory centres in the medulla of the cat. Same strength of stimulation used in both records, viz. 8 volts at 240 stimuli per second. Duration of stimulation shown by signal line. Time, 6 second intervals. Inspiration down, expiration up. A. Maximal inspiratory standstill on stimulation of inspiratory centre. B. Maximal expiratory standstill on stimulation of expiratory centre. (PITTS, MAGOUN and RANSON. *Amer. J. Physiol.*, 1939, **126**, 673.)

The relative positions of these centres are shown in Fig. 504, and their effects on respiration in Fig. 505. Pitts § regards the central mechanism as four-fold, viz. (1) The motoneurone pool in the reticular formation of the medulla; (2) the vagal inhibitory system; (3) the brain-stem inhibitory system; and (4) other inhibitory and excitatory systems: the second and third are essential for normal functioning of the first: the third has been called by Lumsden the "pneumotaxic centre."

---

* *v.* BERNTHAL. *Ann. Rev. Physiol.*, 1944, **6**, 155.

† V. E. HENDERSON and SWEET. *Amer. J. Physiol.*, 1929, **91**, 94.

‡ LUMSDEN. *J. Physiol.*, 1923, **57**, 153, 354; **58**, 81, 111.

§ PITTS. *J. Neurophysiol.*, 1942, **5**, 403.

From the medullary centres on each side, the efferent impulses to the motor nuclei of the respiratory muscles pass down in the ventral and ventro-lateral columns of the cord. Hemisection of the upper cervical cord, *e.g.* on the right side, causes cessation of the contractions of the diaphragm on the same side. There must, however, be commissural fibres joining the motor nuclei on the two sides. If the right phrenic nerve be divided, after hemisection on the left side, the left half of the diaphragm at once commences to contract rhythmically with each respiration.* It is evident that the cessation of respiration after section of the cord is due to the fact that the spinal motor centres cannot act without higher control, and is not due to a condition of shock of the lower spinal centres, since it is possible for impulses to pass down the cord and to cross over to the contra-lateral diaphragm nucleus *immediately* after hemisection of the cord on the side of the nucleus.

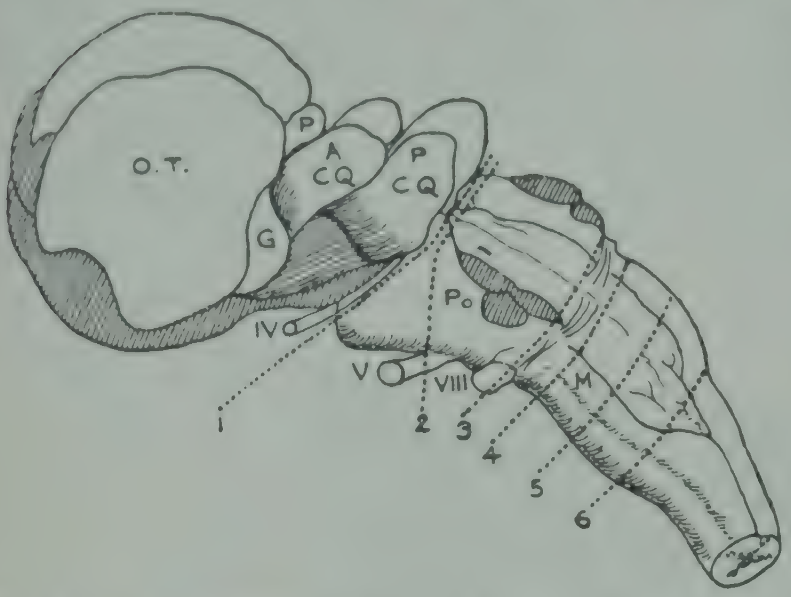


Fig. 504. Diagram of Brain Stem of Cat, showing Levels at which cross section alters the Type of Respiratory Movement. (Lumsden.)

Section 1 leaves respiration unaltered.

Section between 2 and 3, causes development of inspiratory type of respiration (apneusis), if vagi are also cut.

Section 4 causes gasping respiration.

Section 6 abolishes respiratory movements.

Cortical Control. It is common knowledge that respiratory movements can voluntarily be controlled in any direction for a time, and in accordance with

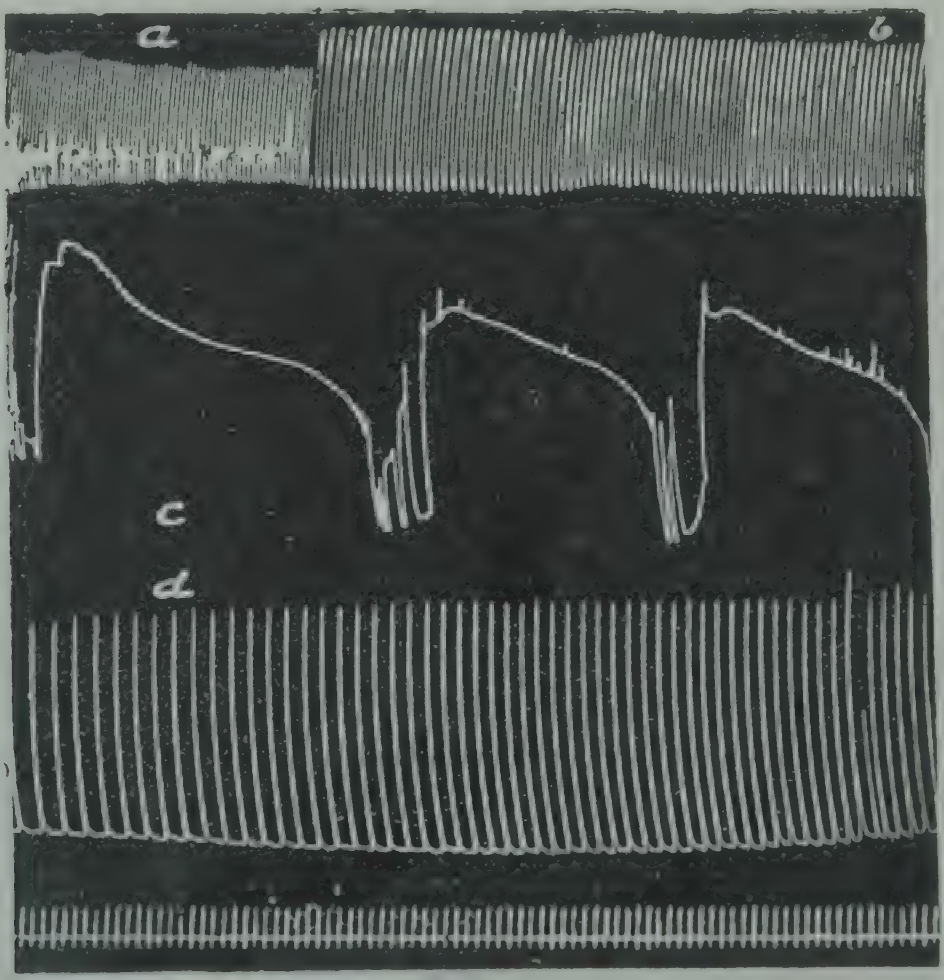


Fig. 505. Tracings of various Types of Respiration in the Cat. (Lumsden.)

(*a*) Normal. (*b*) After vagotomy. (*c*) Prolonged inspiratory tonus (*apneusis*) after section 2, Fig. 504, and vagotomy. (*d*) Gasping respiration after section 4. Time tracing = 5 secs. Inspiration upwards.

---

* Porter. *J. Physiol.*, 1895, **17**, 455.

this, certain parts of the cerebral cortex can be shown to affect the breathing when stimulated. Certain parts of area 6 of the cortex in cats, dogs and monkeys when stimulated cause acceleration, while the excitation of other areas gives inhibition of breathing. These areas have corresponding cytoarchitecture in all these species.*

The Automaticity of the Respiratory Centres. The respiratory centres are often described as automatic. By this we mean that, although they can be affected reflexly, or by chemical changes in the blood, their rhythmic activity is primarily dependent on the metabolism of their own cells.† It is true that the centres continue to function rhythmically after almost complete isolation, *e.g.* after afferent impulses have been cut off from them by section of the brain stem at the corpora quadrigemina, of the cord at the eleventh thoracic segment,

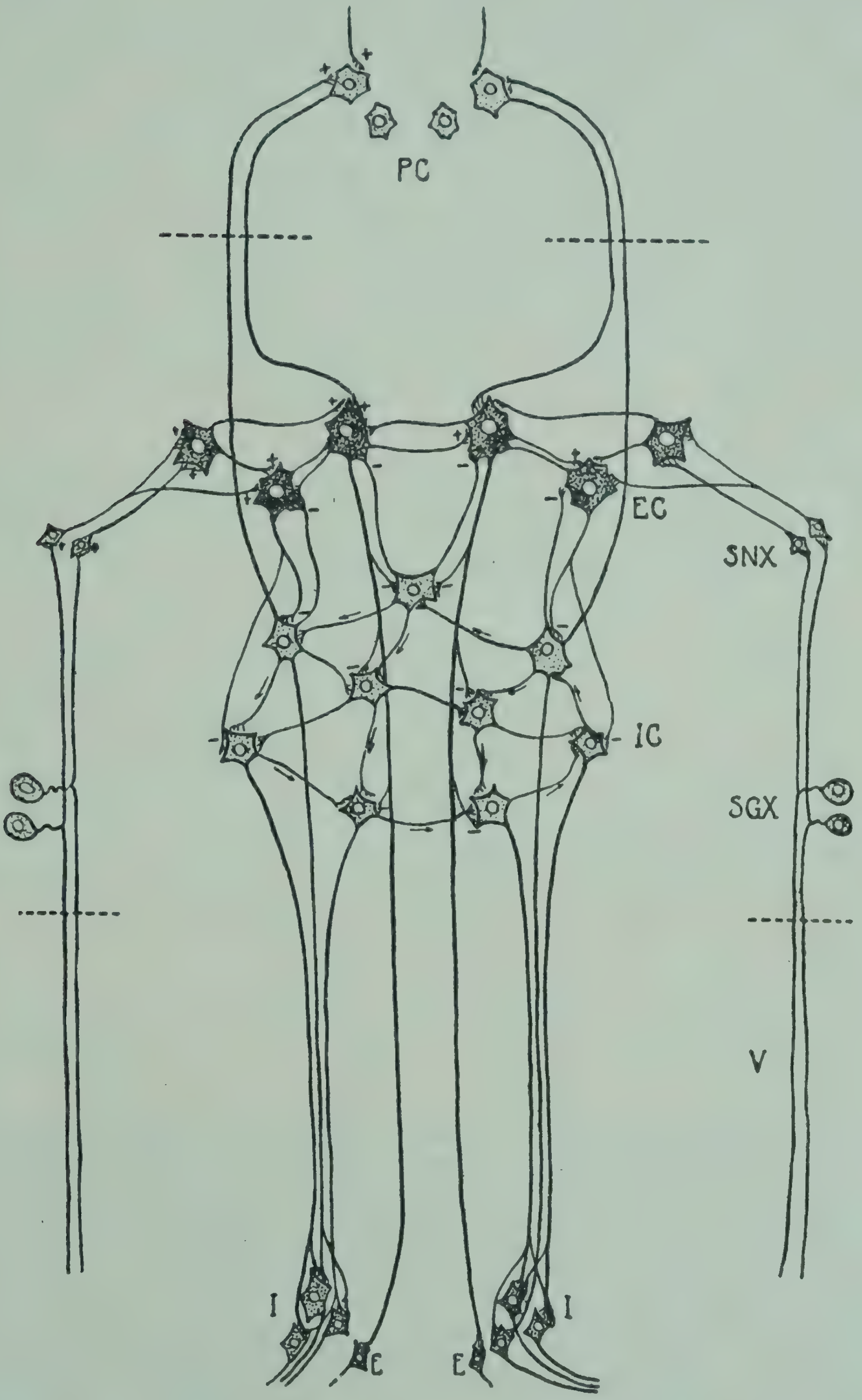


Fig. 506. Diagrammatic representation of possible linkages which would account for rhythmic respiration and apneusis. E, Expiratory, and I, inspiratory motoneurons in cord ; E.C. and I.C., expiratory and inspiratory centres ; P.C. pneumotaxic centre in pons ; V, vagus ; SGX, sensory ganglion of vagus ; SNX, sensory nucleus of vagus. (Pitts, Magoun and Ranson. *Amer. J. Physiol.*, 1939, **126,** 689.)

---

* Smith. *J. Neurophysiol.*, 1938, I., 55.

† Gessell and Moyer. *Q. J. Exp. Physiol.*, 1935, **25,** 13.

of the dorsal roots of all the spinal nerves above that, and of both phrenics, th vagi, fifth, seventh and eighth cranial nerves.* Taking the brain stem centre as a whole, therefore, it does seem that they can act independently of th receipt of afferent impulses. But if we separate off the pontine pneumotaxi centre from the medullary centres by a section across the lower pons, afte section of the vagi, rhythmic respiration is at once replaced by apneusis. Eithe the vagus, or the pneumotaxic centre, must therefore be left in connection wit the medullary centres if these are to continue to function rhythmically. Otherwise the inspiratory centre enters into sustained activity, causing apneusis.

The vagus is connected chiefly with the expiratory centre, and Fig. 50 suggests the kind of linkages in the central nervous system which would accoun for the phenomena of rhythmic breathing and apneusis. The pneumotaxi centre is in the tegmentum of the extreme upper part of the pons, and is connected with the medullary centres by a pathway ventral and slightly medial t the descending root of the fifth cranial nerve.†

The respiratory centres as a whole tend to respond to all stimuli, continuous or rhythmic, by means of rhythmic discharges. This means that the inspiratory and expiratory centres tend to function alternately, as also do many other opposed motor centres, such as those concerned with flexion and extension movements of the limbs.

If the impulses passing down phrenic nerve fibres are recorded, it is seen that at each inspiration a rapid shower of impulses is propagated. This probably means that the inspiratory neurones have reached such a level of depolarisation (p. 247) that repetitive firing ensues, after which the depolarisation subsides. In states of apneusis the discharge is continuous ; both cease on afferent vagal stimulation.‡ As each inspiration proceeds it gathers strength in two ways, viz. by the recruitment of new motor units and by an increasing frequency of nerve impulses in the active units.

By warming the blood in the carotid arteries, there may be a considerable quickening of respiration unaccompanied by any deepening, a condition which is spoken of as *tachypnœa*. On the other hand, we may slow the respiratory movements by cooling the floor of the fourth ventricle.

For Reference

Cordier and Heymans (1935). *Le Centre Respiratoire.* (Paris, Hermann and Cie.)

Gesell (1940). A Neurophysiological Interpretation of the Respiratory Act. *Ergebn. d. Physiol.*, **43**, 477.

Pitts (1946). Organisation of the Respiratory Centre. *Physiol. Rev.*, **26**, 609.

THE CHEMICAL REGULATION OF THE RESPIRATORY MOVEMENTS

Asphyxia. If the proper aeration of the blood be interfered with in any way, the respiratory movements increase in depth and frequency, and if the disturbing factor be not removed, the animal dies of "asphyxia." The asphyxia may be caused by occlusion of the trachea, paralysis of respiratory muscles, by pneumothorax, etc.

The phenomena of asphyxia may be divided into three stages :

(1) In the first stage, that of *hyperpnœa*, the respiratory movements are increased in amplitude and in rhythm. This increase at first affects both inspiratory and expiratory muscles. Then the breathing becomes *dyspnœic.* Gradually, the force of the expiratory movements becomes increased out of all proportion to the inspiratory, and the first stage merges into :

(2) the second, which consists of *expiratory convulsions*, in which almost every muscle of the body may be involved. Just at the end of the first stage

---

* Stella. *J. Physiol.*, 1938, **93**, 10.

† Pitts, Magoun and Ranson. *Amer. J. Physiol.*, 1939, **127**, 654.

‡ Pitts. *J. Neurophysiol.*, 1942, **5**, 403.

consciousness is lost, and almost immediately afterwards there appear a number of phenomena of central excitation extending to almost all the functions of the body. Thus the vasomotor centre is excited, causing widespread vascular constriction. There is often also secretion of saliva, inhibition or increase of intestinal movements, constriction of the pupil, and so on.

(3) At the end of the second minute the expiratory convulsions cease almost suddenly, and give way to *slow deep inspirations.* With each inspiratory spasm, the animal stretches itself out and opens its mouth widely, as if gasping for breath. The whole stage is one of exhaustion : the pupils dilate widely, and all reflexes are abolished. The pauses between the gasps become longer and longer, until, at the end of four or five minutes, the animal takes its last breath.

In asphyxia, two factors are present, namely, an increase of carbon dioxide and a diminution of oxygen in the blood. Each factor is responsible for certain effects, and it is important to distinguish between the separable states of *oxygen deficiency* (local or general), *hypercapnia* (or hypercarbia, *i.e.* CO_2 excess), and *asphyxia*, in which last the two preceding are combined, and to which both factors contribute, though in different ways (*v.* p. 769), just as we saw that both factors are involved in causing the changes in the circulation in asphyxia (p. 651).

The Normal Chemical Stimulus. At various times the stimulus for respiration has been ascribed, either to the increased tension of carbon dioxide, or to the diminished tension of oxygen, in the arterial blood. Increase of this stimulus, it was believed, augmented the activity of the centre and caused hyperpnœa, reduction lowered its activity, so that breathing was suspended—a condition called *apnœa.* Haldane and his pupils believed that the respiratory centre could itself be excited either by excess of carbon dioxide or by lack of oxygen, but that its sensitivity to carbon dioxide was by far the more important factor in the determination of the increased respiratory movements in asphyxia, and was the only chemical variable in the blood which could be regarded as playing a large part in the regulation of the respiratory movements under normal conditions. This factor is well brought out if we investigate in man the effect on the respiratory movements of altering the tensions of the two gases in the air breathed.* If, by this means, we alter the tension of the two gases in the alveolar air, we may assume that the tensions of the gases in the arterial blood are similarly changed.

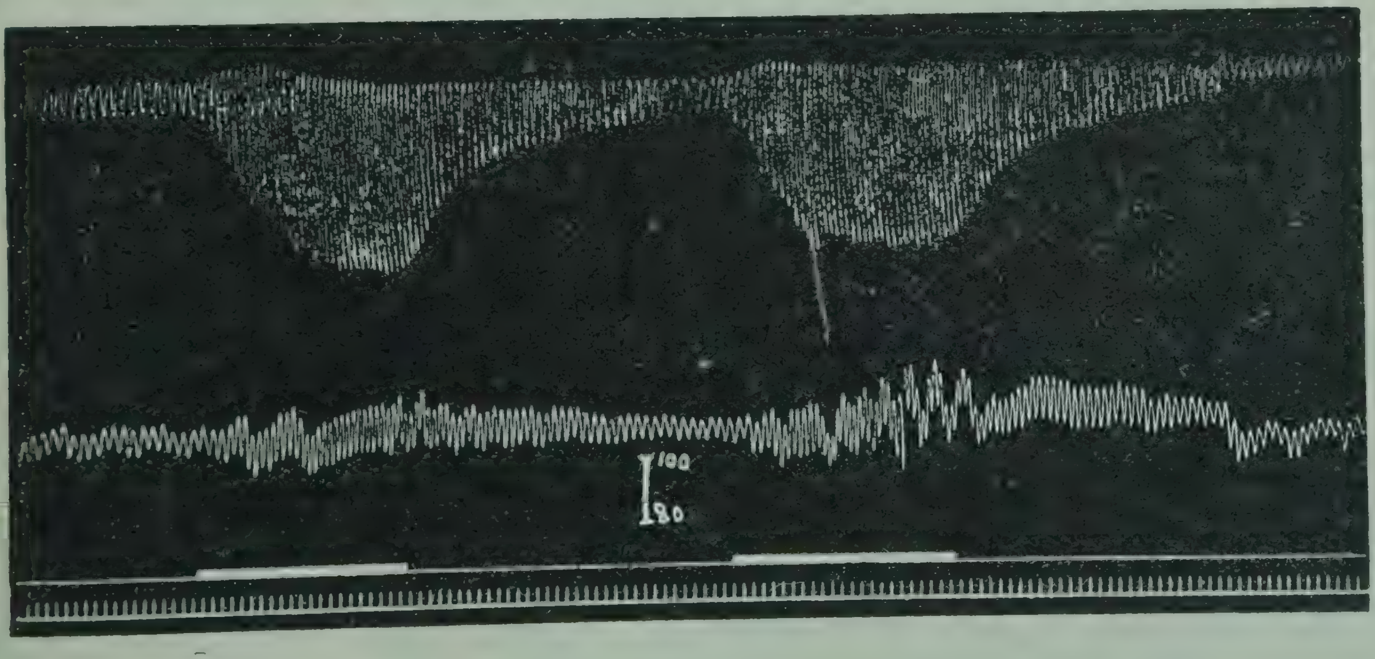


FIG. 507. Effect of CO_2 on Respiratory Movements of Rabbit. (SCOTT. *J. Physiol.*, 1908, **37**, 301.)

Upper line, tracing of diaphragm slip (Head's method). Lower tracing, carotid pressure. During the first period indicated on the signal line the animal breathed 9·6 per cent. CO_2 in air and during the second period 10 per cent. CO_2 with 33 per cent. oxygen. Time tracing = 2 secs. Scale = mm. Hg blood pressure.

---

* HALDANE and PRIESTLEY. *J. Physiol.*, 1905, **32**, 225.

Carbon Dioxide. Even a slight increase in the percentage of carbon dioxide in the air causes an increase, first in the depth and later on in the rhythm, of respiration (Fig. 507). This is shown in the following Table (Haldane and Priestley), which represents the average ventilation effected by the respirations when the subject was breathing either normal air or air containing varying percentages of carbon dioxide. For instance, a rise of carbon dioxide *in the atmosphere* to 1·74 per cent. increases the total alveolar ventilation by 43 per cent.

CO_2 p.c. in inspired air	Mean CO_2 p.c. in alveolar air	Relative rates of alveolar ventilation
0·03	5·71	100
1·74	5·65	143
3.98	6·03	277
5·28	6·55	447

The second column of figures in this Table shows that, in spite of the very large variations in the air breathed, the alveolar content of CO_2 remained practically constant until the CO_2 in the atmosphere was increased to such an extent that the processes of compensation were no longer feasible. We must conclude, therefore, that the pulmonary ventilation is so arranged as to react to the slightest increase of CO_2 tension in the blood, any increase in this giving at once a compensatory increase in ventilation, so that the alveolar CO_2 pressure may be maintained almost constant.

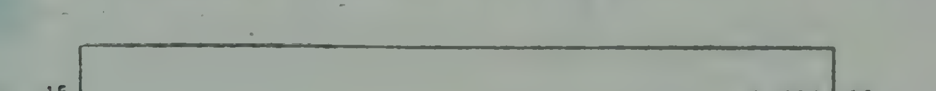


Fig. 508. Effects of Alterations in the Barometric Pressure on the Alveolar CO_2 Tension, the Alveolar CO_2 Percentage and on the Alveolar O_2 Tension. Note that the excitant effects of O_2 lack are not seen until the atmospheric pressure falls below 500 mm. Hg. (Boycott and Haldane. *J. Physiol.*, 1908, **37**, 355.)

That it is the *partial pressure*, and not the percentage amount of CO_2 in the alveolar air, and therefore in the blood, which is really the determining factor, is shown by a comparison of the composition of the alveolar airs of subjects under different atmospheric pressures. Thus at sea level (barometer = 760 mm.) the alveolar CO_2 content of a subject was 5·3 per cent., which represented an alveolar CO_2 pressure of 37·8 mm. Hg. When the subject was placed in an air-chamber compressed to a pressure of 4,640 mm., the mean percentage of CO_2 in the alveolar air was 0·9, corresponding, however, to a *tension* of 41·3 mm. Hg. Thus, the pressure of CO_2 in alveolar air remains practically constant with widely varying limits of atmospheric pressure as well as with very different percentages of CO_2 in the inspired air, showing that the reactions of the organism are directed so as to maintain, by alterations in the ventilation, a constant *tension* of this gas in the alveoli, and therefore in the arterial blood (Fig. 508).

Actually it is the ratio of CO_2 to $NaHCO_3$ in the blood (*i.e.* the pH of the blood) which is kept constant. Hence if the bicarbonate of the blood is altered, the CO_2 of the blood is adjusted by an alteration of the breathing. Just after a meal, for instance, when the HCl of the gastric juice is being separated off from the blood, there remains an excess of base, in the form of bicarbonate in the blood. This leads to a reduction of the breathing so that the alveolar CO_2 shows a slight rise.* A similar result follows if bicarbonate is taken by the mouth † and the converse result if acid or ammonium chloride is taken.‡ In the latter case the ammonia is converted into urea leaving the HCl behind in the blood. The alveolar CO_2 shows a definite fall; it does this under any circumstances, *e.g.* in diabetes, where there is an excess of acid in the blood. In all cases the effect of the response is to regulate the output of CO_2 so that the ratio of $\frac{CO_2}{NaHCO_3}$ remains constant.

While these facts and deductions are incontrovertible, they do not justify the conclusion that the effects of altered CO_2 content of the blood are exerted directly upon the respiratory centres. The effect might be upon some peripheral structure from which the respiratory centre is affected secondarily by reflex action. Experiments on man do not permit of an answer to this objection, which, as will be shown later, is in part justified. We may state here, however, that even when, in animals, all likely afferent nerves are severed, respiration is still augmented by a rise in the CO_2 content of the blood supplied to the centre, so we may conclude that the centre itself does respond directly to that chemical change in the blood; it responds less than in the intact animal, however, as the effect is partly reflex.

STIMULATION BY OXYGEN LACK. Very different are the phenomena observed on sudden alteration of the partial pressure of oxygen (Figs. 508 and 509). Here, within wide limits, the partial pressure of oxygen in the alveolar air is determined by its pressure in the inspired air. Thus, if we take the above series of observations, with a pressure of 646 mm., the *percentage* of oxygen in the alveolar air was 13·19, corresponding to a *tension* of 78·6 mm. Hg. At an atmospheric pressure of 755 mm. the percentage of oxygen in the alveolar air was 13·97, corresponding to a tension of 98·5 mm., which we may take as the

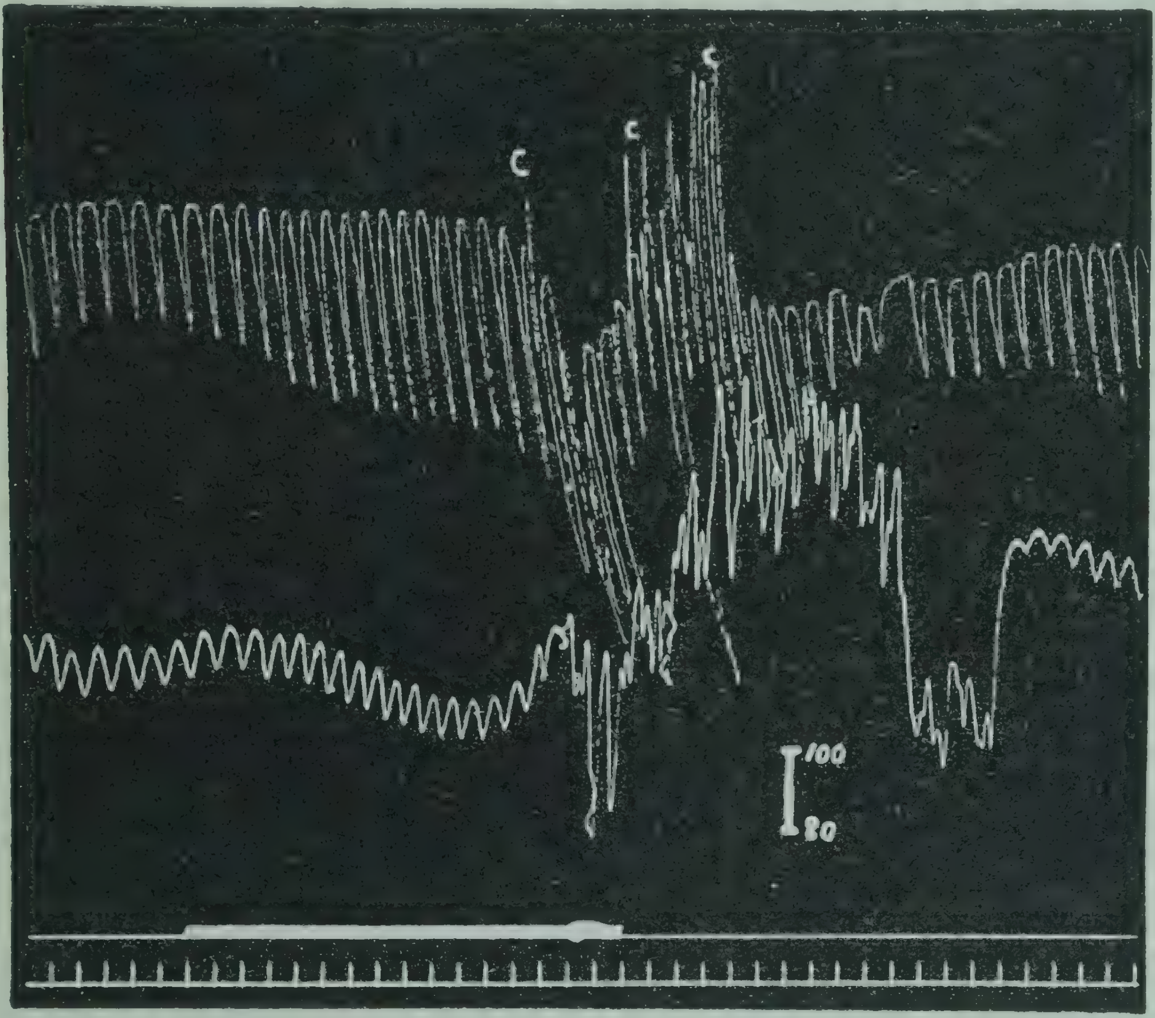


FIG. 509. Effects of Oxygen Lack. (SCOTT.)

Upper tracing, diaphragm slip; lower tracing, carotid blood pressure. During time indicated by signal, 5 per cent. oxygen in nitrogen was inhaled. c = convulsion.

---

* DODDS. *J. Physiol.*, 1921, **54**, 342; DODDS and BENNETT. *J. Physiol.*, 1921, **55**, 381; BENNETT and DODDS. *Brit. J. Exp. Path.*, 1921, **2**, 58.

† DAVIES, HALDANE and KENNAWAY. *J. Physiol.*, 1920, **54**, 32.

‡ J. B. S. HALDANE. *J. Physiol.*, 1921, **55**, 265.

normal figure at the sea-level. In air compressed to a pressure of 1,261 mm. the percentage of oxygen was 16·79, corresponding to an alveolar oxygen tension of 203 mm. Hg.

Similar results are obtained by altering the percentage of oxygen in the air breathed at constant pressure. The oxygen tension or percentage in the inspired air can be lowered from its normal of 20·95 to 13 per cent., at which the alveolar oxygen is about 8 per cent. (= 57 mm. Hg), without altering the depth or rhythm of respiration, and, in fact, without any change being noticed by the subject of the experiment. With a further reduction of the oxygen content, however, there is increased pulmonary ventilation (Fig. 509), but the diminution in oxygen may be pushed to such an extent that the patient becomes blue, without any considerable distress being caused. In fact, the subject may suddenly lose consciousness before he has felt serious discomfort. In all probability such effect as lack of oxygen has upon respiration is not exerted directly upon the respiratory centre, as will be shown later.

The difference in the sensitiveness of the response to increase of carbon dioxide and lack of oxygen respectively, is well shown by an experiment of Haldane's, in which the same person breathed in and out of a bag, in the first place allowing the carbon dioxide produced in respiration to accumulate, and in the second, removing the carbon dioxide by means of soda lime, so that the sole effect of respiration was to produce a continual diminution in the percentage of oxygen. In the first case, when the carbon dioxide was allowed to accumulate, it was found that intolerable hyperpnœa was produced when the content of the bag had reached 5·6 per cent. carbon dioxide and 14·8 per cent. oxygen. When the carbon dioxide was absorbed, no hyperpnœa was produced, and the experiment was stopped as soon as the subject was becoming blue in the face and experienced slight throbbing in the head. The pulse frequency had gone up from 80 to 108. The bag contained no carbon dioxide and only 8·7 per cent. oxygen.

We must conclude that the respiratory apparatus is peculiarly sensitive to variations in the pressure of carbon dioxide. Since the animal has developed a mechanism by which the reaction of the blood can be rapidly adjusted by varying the excretion of carbon dioxide, whilst the excretion of other acids is relatively slow, carbon dioxide may be regarded as, in effect, the normal respiratory stimulant. Although the breathing can be excited by lack of oxygen, this effect is rarely brought into play. Under all ordinary circumstances, an increased need for oxygen is associated with an increased accumulation of carbon dioxide in the body, and the augmentation of respiration, produced by the excitatory effect of a small excess of carbon dioxide tension in the blood, suffices also to provide fully for the increased needs for oxygen.

THE REFLEX REGULATION OF RESPIRATION

Afferent Nerves. Although the sensibility of the respiratory centre to CO_2 is the most important factor in determining the respiratory ventilation, the respiratory centre is modified in a large degree by impulses arriving at the centre along the vagi and the carotid sinus nerves. Through other sensory nerves the respiratory movements can at times be altered reflexly, but under normal circumstances it is only through these nerves that a continuous stream of impulses passes to the centre, so that every respiratory movement is modified by these impulses.

If both vagi be divided, a change in the respiratory rhythm is at once seen. The movements become less frequent, but are increased in amplitude. If, now, the central end of one of the vagi be stimulated with weak stimuli at a low frequency, *e.g.* 30 per sec., an inspiratory increase of the movements of respiration results. With more rapid stimuli, *e.g.* 250 per sec., an expiratory hypertonus results, or a definite expiration. The vagus is connected mainly with the expiratory centre, excitation of which inhibits inspiration and provokes expiration.

Expiration may be strikingly brought about by stimulation of the central end of the superior laryngeal nerve, which produces first an inhibition of inspiration and then a forcible expiration. The superior laryngeal nerve supplies sensory fibres to the mucous membrane of the glottis, and we know that the slightest irritation of these fibres—the presence of a crumb or a particle of mucus—causes forcible expiratory spasms, with spasmodic closure, of the glottis, which we term a cough. It must not be imagined that the fibres of the superior laryngeal nerves are concerned in the reflex maintenance of the normal respiratory rhythm. They are cited here merely because the result of their stimulation somewhat resembles that which would be caused by stimulation of the analogous expiratory fibres which run in the trunk of the vagus from the *lungs* to the respiratory centre.

Since section of both vagi causes slowing of respiration, impulses which exert some influence to quicken respiration must travel up the vagi from the lungs. The respiratory movements cause an alternate inflation and deflation of the lungs, and it appears that these changes in the volume of the lungs start or stop the accelerating impulses that travel up the vagus nerves. To test the truth of this hypothesis, it is necessary to see the result, on the respiratory movements, first of distension, and secondly of collapse, of the lungs.

When it is wished to study the effects of artificial distension or collapse of the lungs we may use the method described by Head.* In the rabbit, a slip of the diaphragm serves as a sample of the diaphragm, the contractions of which may be recorded, uninfluenced by passive movements. Alternatively, in the dog, the movements of the larynx up and down with each respiratory movement can be recorded.

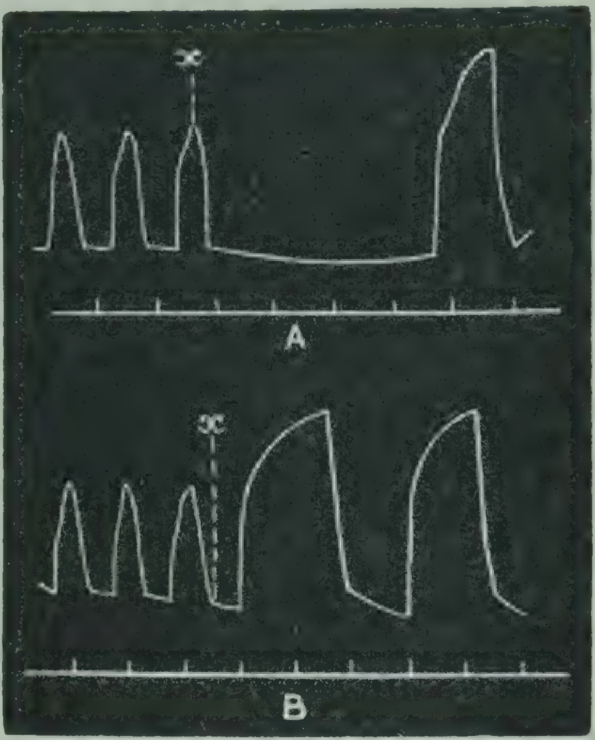


FIG. 510. Effects of Distension and Collapse of Lung. Contraction of diaphragm slip of rabbit. Upward movement = inspiration (contraction). (FOSTER.)

In A, the trachea is closed at x, the height of inspiration; a long pause follows, and finally an inspiration (a very powerful one) sets in.

In B, the trachea is closed at the end of expiration, x; powerful inspirations follow at once.

The effects of simply closing the trachea at the end of inspiration or of expiration are shown in Fig. 510, and similar results are obtained by experiments on man.† A still more marked effect is produced if the lungs be artificially inflated. The inflation produces an instantaneous and complete relaxation of the diaphragm (Fig. 511), which by clamping the tracheal tube may be prolonged for several seconds. Forcible deflation of the lungs produces the opposite effect, viz. contraction of the inspiratory muscles.‡

When the action potentials in a single vagus fibre from the lungs are recorded, it is found that it conveys impulses in rapid succession each time the lungs are

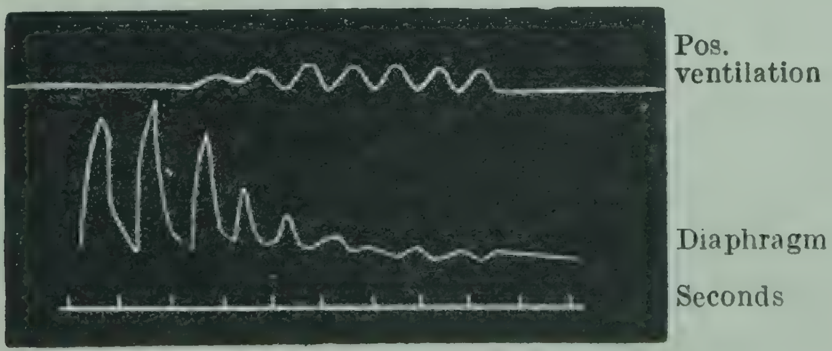


FIG. 511. Positive Ventilation. (HEAD.)

Under the influence of positive ventilation, the inspiratory contractions of the diaphragm become less and less till they disappear completely.

---

* HEAD. *Journ. Physiol.*, 1889, **10**, 1.

† HALDANE and MAVROGORDATO. *J. Physiol.*, 1916, **50**, Proc. xli; LAMB. "An Introduction to Human Experimental Physiology," p. 183, 1930. (Longmans.)

‡ HEAD. *J. Physiol.*, 1889, **10**, 1; CREED and HERTZ. *Ibid.*, 1933, **78**, 85.

distended, the impulses being more numerous and in more rapid sequence the more powerful the distension.* The lung evidently possesses stretch-receptors similar to those of muscle, and like them causing reflexly an inhibition or moderation of the movement that produced the stretch—in this case the act of inspiration (Fig. 512). In deflation of the lung these impulses cease; but on forcible suction fresh impulses ascend the vagus. These are probably not present in normal breathing movements.

It is not known where the receptors which respond to inflation of the lungs are situated, though it seems probable that they are in the walls of the bronchi and bronchial tubes. Fibres of the vagus have been traced to the bronchialis muscle, where they terminate in sensory endings resembling those of the carotid sinus, and presumably, like these, acting as stretch receptors.†

The Hering-Breuer Reflex. These inhibitory and augmentor effects of changes in the volume of the lung must also result from the normal movements of respiration. Let us consider, for instance, what will happen if the influence of the two vagi could be suddenly restored after these nerves have been divided. (This experiment can in fact be realised if the functional division of the vagi be effected by cooling.) The animal would be breathing slowly and deeply. If, at the beginning of an inspiration, the vagi became functional, the expansion of the lungs caused by the inspiratory movement would send impulses up the vagus to the expiratory centre, which would arrest the movement of inspiration. The movement of expiration would then begin, but the collapse of the lungs thereby produced would stop the impulses up the vagi, and so allow the inspiratory centre to cause an inspiratory movement. Both inspiration and expiration would therefore be shortened, and the successive movements would follow one another at a shorter interval than if the vagi were not functional. This reflex is called the Hering-Breuer reflex.‡ It is abolished when both vagi are divided (Fig. 513). By its operation the *rhythm* of the respiratory centre must be determined reflexly through the agency of the vagi, while the chief factor in determining the total pulmonary *ventilation* is, as we have seen, the carbon dioxide tension of the blood.

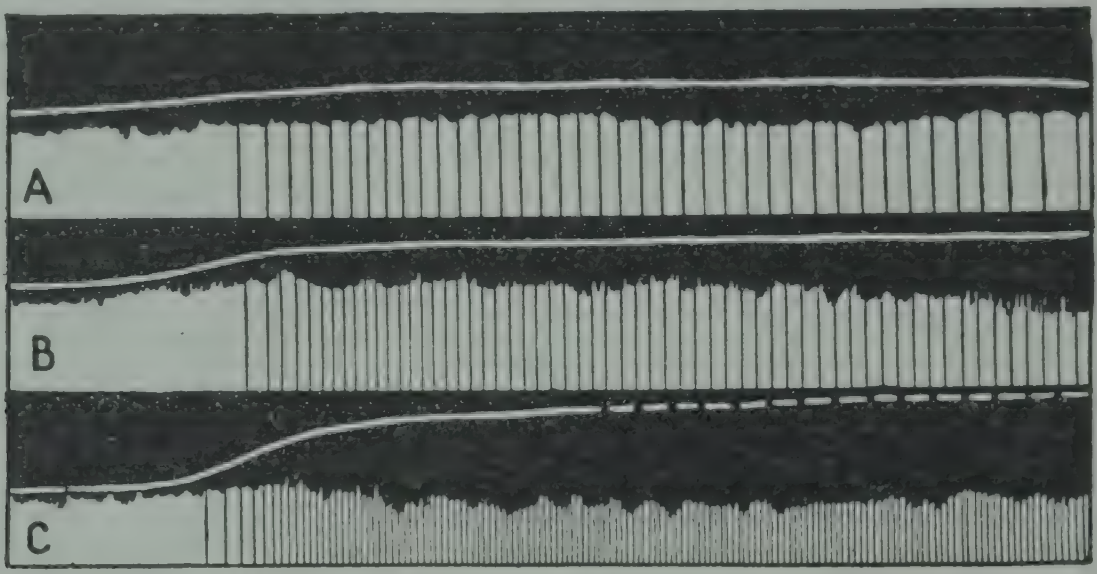


FIG. 512. Action potentials in pulmonary vagus fibres of cat.

A. Spinal Cat. Inflation of 65 c.c. . Max. frequency = 80 per sec.
B. Spinal Cat. Inflation of 115 c.c. . ,, ,, = 120 ,,
C. Decerebrate Cat. Natural breathing (one breath).

The light line shows degree of inflation of lung. (ADRIAN. *J. Physiol.*, 1933, **79**, 332.)

Though the effect of double vagotomy is to produce slowing of the respiration, it is necessary to answer the objection that this may not be due to the section of the pulmonary fibres in the vagus, but to the division of fibres running in the trunk of the vagus from some other source. We know that other afferent fibres which would also be divided can influence respiration, *e.g.* the fibres of the "aortic" nerve, from the cardio-aortic area, and perhaps the fibres of the recurrent laryngeal nerve.* These objections have been answered by the experiment of dividing these nerves, while leaving the pulmonary branches of the vagus intact, a procedure which leaves the respiratory rate unchanged, and then by the converse experiment of sectioning only the pulmonary branches of the vagus, which produces the same slowing as division of the entire cervical vagus,† and, like it, abolishes the Hering-Breuer reflex (Fig. 513).

---

* ADRIAN. *J. Physiol.*, 1933, **79**, 332.

† GAYLOR. *Brain*, 1934, **57**, 143.

‡ HERING and BREUER. *Sitz. d. Wien. Akad. Math.-Nat. Cl.* (2), 1868, **57**, 672; **58**, 909.

Although extreme collapse of the lung may also cause afferent impulses to pass up the vagi, it seems certain that these are not engendered by the moderate degree of collapse accompanying normal expiration. When such impulses are produced by the unusual stimulus of strong suction, they seem able to excite the opposite reflex to that produced by inflation, viz. an inspiratory movement. The production of inspiration by weak slow stimulation of the central end of the vagus is further evidence that such a reflex is feasible though it does not justify the assumption that it is involved in the production of normal inspiratory movements.

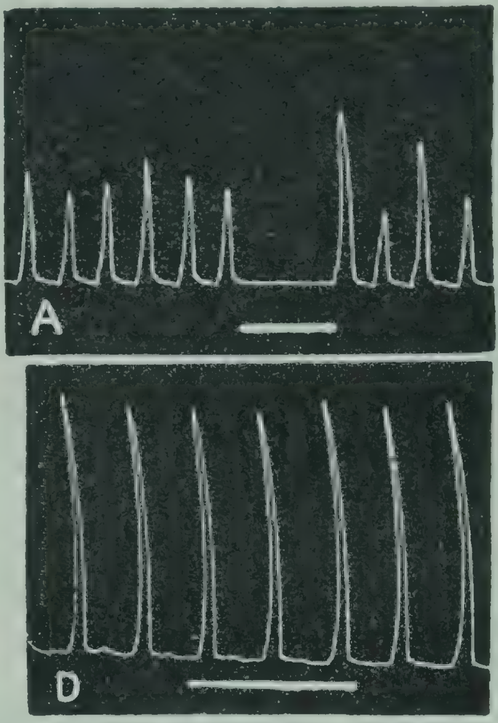


FIG. 513. Respiratory movements of larynx of chloralosed dog. A. The left vagus has been divided in the neck. Inflation of lungs during white line arrests respiration. D. The pulmonary branches of the right vagus have now been divided. The Hering-Breuer reflex is abolished. (ANREP and SAMAAN. *J. Physiol.*, 1932, **77**, 1.)

Collaboration of Nervous and Chemical Control. After division of both vagi, the total pulmonary ventilation does not, as a rule, undergo any marked change, and the aeration of the blood is carried out as well as in the intact animal. The effect of the impulses travelling up the vagi, by reining in the activity of the centre, is to economise its energy and the energy of the respiratory muscles. The result of the vagal impulses will therefore be to increase the delicacy of the respiratory centre and make it more susceptible to slight changes in the carbon dioxide tension of the blood, while maintaining a sufficient margin of energy to meet the increased needs thrown on the respiratory mechanism by augmented metabolism, such as occurs in violent muscular exercise. The importance of the vagus action for the organism is shown if we put an increased strain on the respiratory mechanism, as, for instance, by increasing the carbon dioxide in the air breathed.‡ In the intact animal, this

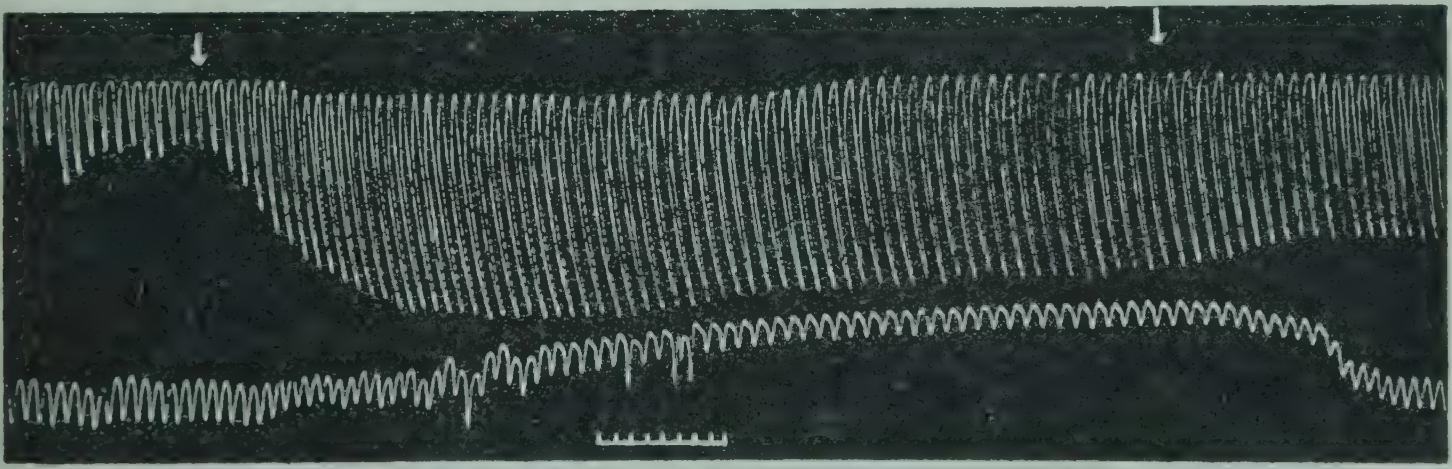


FIG. 514. Effect of 10·6 per cent. CO_2 in a Mixture containing 23·3 per cent. O_2 on a Rabbit with both Vagi divided. (F. H. SCOTT. *J. Physiol.*, 1908, **37**, 301.)

The gas was administered between the arrows. Zero line of blood pressure is 32 mm. below bottom of tracing. Compare this figure with Fig. 509, p. 757.

---

* SHARPEY SCHAFER. *Q. J. Exp. Physiol.*, 1919, **12**, 231.

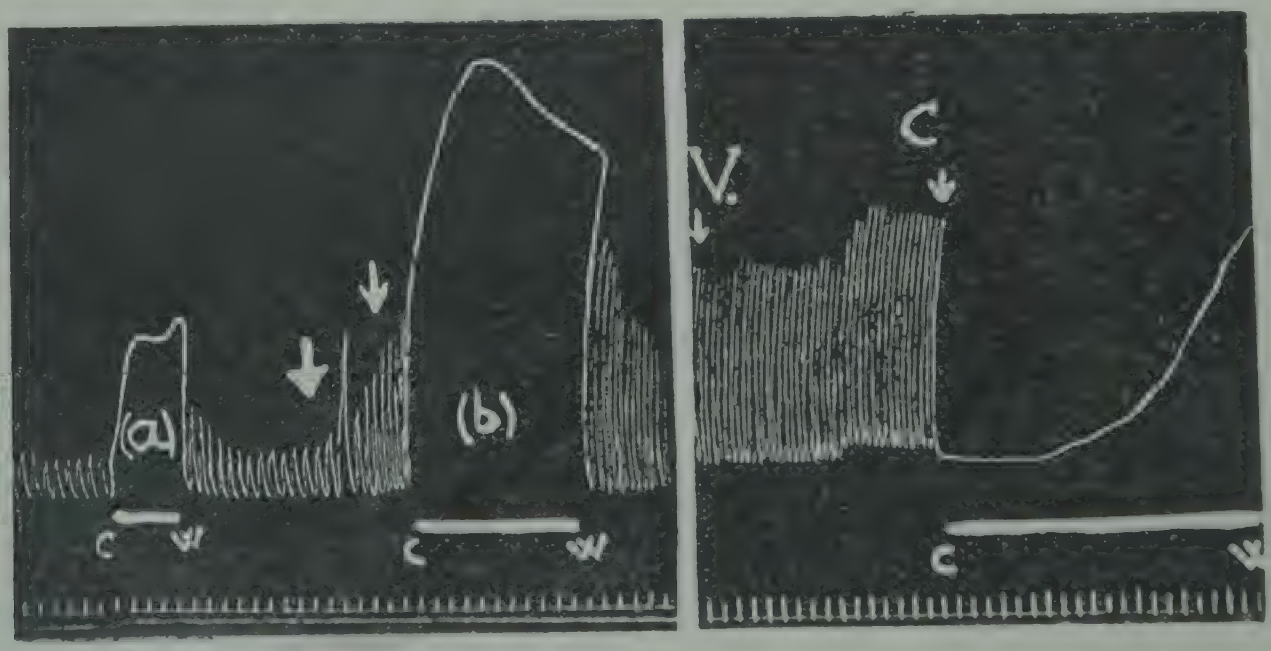
† ANREP and SAMAAN. *J. Physiol.*, 1932, **77**, 1; HAMMOUDA, SAMAAN and WILSON. *J. Physiol.*, 1943, **101**, 446.

‡ SCOTT. *J. Physiol.*, 1908, **37**, 301.

procedure leads first to increased depth, and later to increased frequency of respiration, the total ventilation being thereby augmented to such an extent as to keep the alveolar tension of carbon dioxide almost constant. If the same percentage of carbon dioxide be administered to a chloralosed rabbit after section of both vagi, the effect is deepening of respiration, but not quickening (Fig. 514). Each inspiratory movement, however, was already large, so that the margin by which increase of pulmonary ventilation is possible, by increase of depth of respiration alone, is not so great as in a normal animal. Therefore, since no quickening of respiration takes place, the increased ventilation rapidly becomes inadequate for the maintenance of the normal alveolar carbon dioxide tension. In the following Table the total amounts of pulmonary ventilation, obtained on administration of mixtures containing carbon dioxide to a rabbit before and after section of the vagi, are compared.

	Respirations per minute	Vol. of each respiration	Total ventilation per minute
		c.c.	
Respiration with air . . .	72	19	1,368
,, 4·2 per cent. CO_2 .	96	25	2,400
,, 8·6 per cent. CO_2 .	97	29	2,813
VAGI DIVIDED			
Respiration with air . . .	45	29	1,305
,, 4·2 per cent. CO_2 .	45	34	1,530
,, 8·6 per cent. CO_2 .	42	38	1,596

The Vagus and Apneusis. We have seen that the respiratory centre (*i.e.* the inspiratory and expiratory centres) cannot function rhythmically when isolated from both the pneumotaxic centre and the vagus. Hence, if the influence of the pneumotaxic centre is removed by section across the upper pons, suspension of the action of the vagi, as by local cooling, produces a prolonged holding of the breath in the position of inspiration, a state called apneusis. This is due to a condition of *tonic* activity of the inspiratory centre and is caused by the presence of carbon dioxide in the blood reaching the centre. If the carbon dioxide tension in the blood is raised, as by breathing air containing CO_2, the apneustic contraction of the inspiratory muscles is increased, while if the



A B

FIG. 515. Illustrating the effect of CO_2 content of the blood on apneusis. Cat decerebrated by section through the upper pons. Between c and *w*, vagi were blocked by local cooling. Between arrows in A, inhalation of air containing 8 per cent. CO_2. At V, in B, excessive artificial ventilation of the lungs with air. Time, 4 secs. (STELLA. *Journ. Physiol.*, 1938, **93**, 263.)

CO_2 in the blood is lowered by excessive artificial respiration, the apneusis remains in abeyance until such time as the CO_2 of the blood has mounted up to a level adequate to excite the centre (Fig. 515). When the action of the vagi is restored again by re-warming the vagi, apneusis ceases because the expiratory centre is now stimulated by the afferent impulses along the vagi from the stretched lung, and so rhythmic breathing is restored.

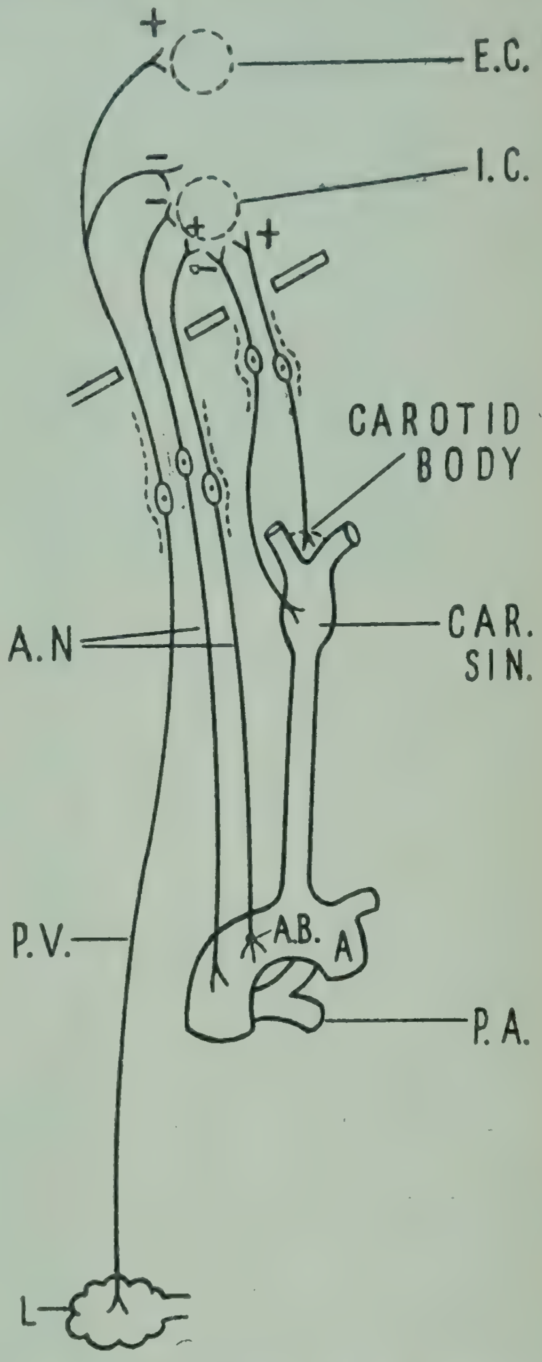


Fig. 516. Relations of vagus, aortic nerve, and nerves of the carotid region to control of respiration. E.C., Expiratory centre; I.C., Inspiratory centre; A.N., Aortic nerve; A.B., Aortic body; A., Aorta; P.V., Pulmonary vagal afferent; L., Lung alveolus.

Chemical Respiratory Effects on the Carotid Body. An important means by which the activity of the respiratory centre may be affected is the influence of the chemical composition of the blood upon nerve endings ("chemoreceptors") in the carotid body and aorta. These we have seen to be important regions in connection with the regulation of arterial blood pressure and heart rate; they, and especially the carotid body, or glomus, must now be regarded as also of importance in the regulation of the respiratory movements.

It has been shown by Heymans and his collaborators,* that when the carotid sinus region (including the glomus), with its nerve supply intact, is isolated from the circulation by appropriate ligatures, and perfused with blood which never enters the general circulation, and so never reaches the respiratory centre at all, definite respiratory reflexes are produced when the composition of that blood is changed; these are of the same nature as would have resulted had the blood perfused through the carotid region been circulating through the entire animal (Fig. 517).

Thus, perfusion of the sinus region with blood containing a deficit of oxygen, or an excess of CO_2, or with asphyxial blood, causes hyperpnœa, while perfusion with oxygenated blood deficient in CO_2 reduces the breathing. Similarly, if the carotid sinus of a dog A is perfused with blood from the carotid artery of a dog B, this blood being returned to B and not allowed to enter the general circulation of A, then asphyxia of B leads to hyperpnœa of A. The carotid body is especially sensitive to oxygen lack and less so to carbon dioxide excess. When the carotid bodies are denervated (and even if the sinus innervation is preserved), these chemical effects produced by alteration of the blood perfusing it are no longer seen. Moreover, the usual effects of oxygen lack or of HCN (which abolishes oxygen usage of tissues) on respiration in the whole animal are not seen if the carotid bodies and aorta are denervated. Under these circumstances oxygen lack rather depresses than excites the respiratory centres.† (Fig. 518).

That the chemical effects arise from the carotid body and not from the stretch-receptors of the carotid sinus has been shown by denervating the latter while leaving the innervation of the former intact.‡ Under these conditions a regular succession of impulses was shown to pass along the carotid body nerve normally, *i.e.* provided the O_2 tension in the arterial blood was below 100 mm. Hg. and the CO_2 pressure over 35 mm. If and when the O_2 tension fell, or CO_2 pressure rose, the frequency of the impulses was augmented. It seems reasonable to suppose, then, that impulses from these chemo-receptors represent one of the normal driving forces that keep the respiratory centre working. In proof of this are the facts that denervation of the carotid *body*, or inhalation of pure oxygen, lead to a small reduction of breathing in animals.

---

* Heymans and Bouckaert. *J. Physiol.*, 1930, **69**, Proc. xiii; Bouckaert, Dautrebande and Heymans. *J. Physiol.*, 1930, **71**, Proc. v.; Heymans, Bouckaert and Dautrebande. *Pflüger's Arch.*, 1932, **230**, 283.

† Selladurai and Wright. *Q. J. Exp. Phys.*, 1932, **22**, 233.

‡ Bogue and Stella. *J. Physiol.*, 1935, **83**, 459; Samaan and Stella. *J. Physiol.*, 1935, **85**, 309; v. Euler *et al.* *Skand. Arch. Physiol.*, 1939, **83**, 132.

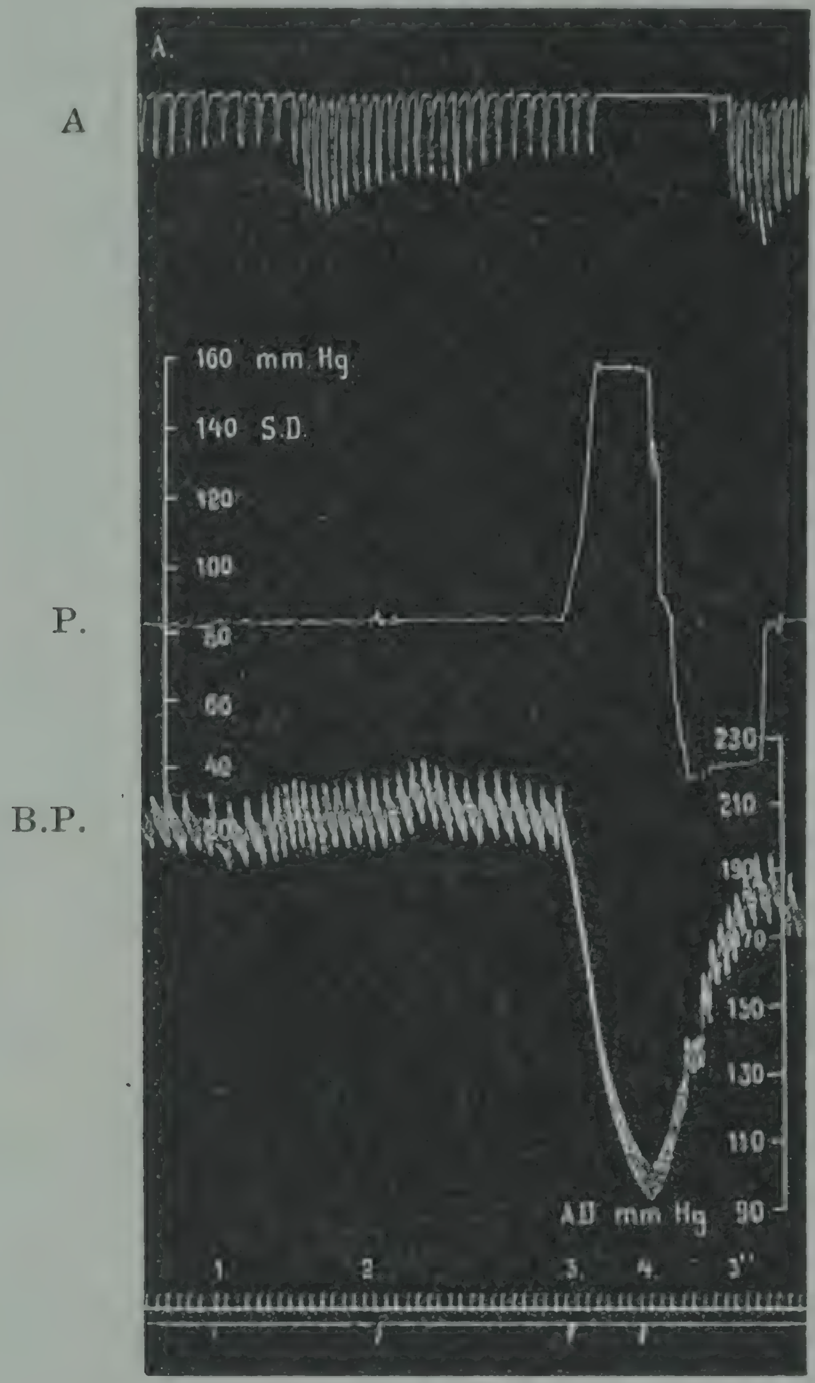


FIG. 517. Effect on respiration of dog of alteration of CO_2 content, and of pressure of fluid in the carotid sinuses. The isolated carotid sinus was perfused with a CO_2-Ringer solution from a pump. At the start the fluid contained 11 vols. CO_2 per cent. and had a pH of 7·25 ; at 1 changed to fluid containing 79 vols. p.c. CO_2 and pH also 7·25 (by increasing the $NaHCO_3$) ; at 2 returned to original fluid ; at 3–4 the perfusion pressure was raised ; at 4 pressure lowered to 40 mm. Hg. A = respiration. P = perfusion pressure. B.P. = arterial pressure. Time = 3 seconds. (HEYMANS, BOUCKAERT and DAUTREBANDE. *Pflüger's Arch.*, 1932, **230,** 283.)

The response to oxygen lack is thus mainly due to a reflex from the carotid and aortic bodies ; to raised CO_2 tensions, the respiratory centre is probably more sensitive than are the chemo-receptors.

That alteration in the composition of the blood in the aorta caused respiratory reflexes was shown by the fact that if the head of a dog was separately perfused, changes of blood composition in the trunk still caused respiratory changes in the larynx movements, provided the vagi were intact.* These reflexes arise from the aortic body, which resembles the carotid body in structure and function.

---

* HEYMANS and HEYMANS. *Arch. int. pharmac. therap.*, 1927, **33,** 273.

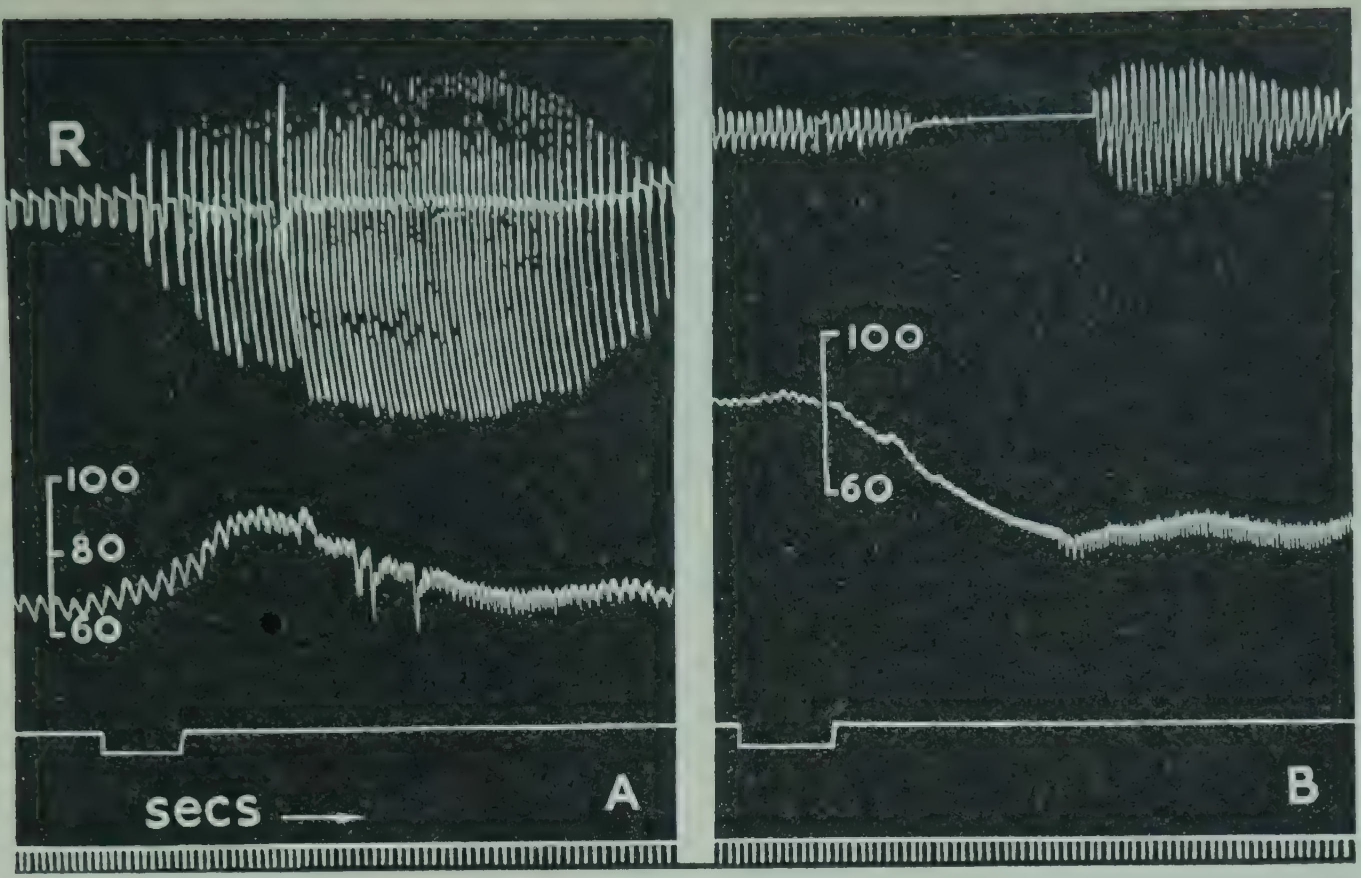


FIG. 518. Rabbit, Nembutal. A. Intravenous injection of 0·35 mg. HCN, vagi and carotid body innervation intact. B. The same dose after section of vagi and denervation of carotid bodies. R. Respirations; below, arterial pressure. Note the short time to produce effect in A (*i.e.* to reach carotid body) as compared with B (to reach medulla). The late augmentation of breathing in B is probably due to CO_2, acting on the centres.

Question of a Single Chemical Influence. Considerable discussion has taken place as to the exact nature of the stimulus to the respiratory system, when the chemical composition of the arterial blood is altered. The breathing is particularly sensitive to changes in blood CO_2. For instance, when the CO_2 pressure of the alveolar air is raised by only 2 mm., the breathing is increased about threefold.* Yet there would be a decrease of pH of the blood of only 0·012. Can we, then, regard the respiratory apparatus (*i.e.* the centre *plus* the peripheral structures such as carotid body, etc.) as exquisitely sensitive to the H-ion concentration of the blood, and consider this as the real stimulus ? The answer is that we cannot, because, under some conditions, the centre behaves as though it were indifferent to the blood reaction, though sensitive to CO_2 or, sometimes, to the HCO_3' ion.

Thus, after administration of ammonium chloride, an acidæmia results, because the NH_4 is converted to urea, leaving HCl behind in the body. Hence there is a rise of H-ion concentration of the blood (*e.g.* a *fall* of pH of, say, 0·07) yet only a small increase of breathing (say 0·5 l./min.). But the same subject when breathing air plus CO_2 to get a smaller fall of pH (*e.g.* 0·04) showed a larger increase of breathing (*e.g.* 10 l./min.).† Again, intravenous injection of $NaHCO_3$ solution raises the pH of the blood, but does not notably depress the breathing. It is admittedly a complex problem, because we cannot easily alter the pH of the blood without also altering the CO_2 pressure in the opposite direction. Hence, as a rule, a fall of pH is accompanied by increased pulmonary ventilation.

The real reason for the excitation produced in the respiratory centre by lack of oxygen, which, as we have seen, is a reflex from the carotid body, is even more obscure. Lack of oxygen might reasonably be expected to act, not as a stimulating, but as a paralysing agent, both on the cells of the respiratory centre and on any peripheral structures. Can we find any common ground on which to explain the action of both oxygen lack and CO_2 excess ?

Freely interpreting the views of Gesell, we might regard the excitation of the cells of the respiratory centre and reflex zones as dependent upon a complex heterogeneous equilibrium between these cells and the blood. According to this view, the cells of the respiratory centre and of the chemo-receptor zones are excited when their cytoplasm becomes more acid. These cells are believed to be continually producing lactic acid, which is normally being removed, chiefly by oxidative processes, but partly by passing into the blood. There would thus be a state of balance between the rate of production and the rate of removal of the acids (lactic and CO_2), and on this balance would depend the reaction, and through that, the excitability, of the nervous structures. Oxygen lack also has a paralytic effect, however, and this is its main action on the cells of the respiratory centre. A lowered oxygen tension in the carotid body would lead to a diminished rate of oxidative removal of lactic acid (as in any other cells), and consequently to its accumulation in the cells, which would cause their H-ion concentration to rise. If the oxygen utilisation of the structures is abruptly checked, as by the injection of a trace of cyanide into the blood, the same thing would happen at once, and there would be an instant hyperpnœa. Addition of lactic acid or carbon dioxide to the blood, with increase of its hydrogen-ion concentration, would also indirectly affect the reaction of the cells of the centre and of the carotid body. In the long run the addition of any strong acid to the blood is equivalent to addition of carbon dioxide, since it displaces that gas from the bicarbonate of the plasma; carbon dioxide diffuses so readily through the walls of all animal cells that it very speedily affects the hydrogen-ion concentration of the cell interior, and it is probably this high penetrating power which explains the dominant *rôle* played by carbon dioxide in regulating the respiration.

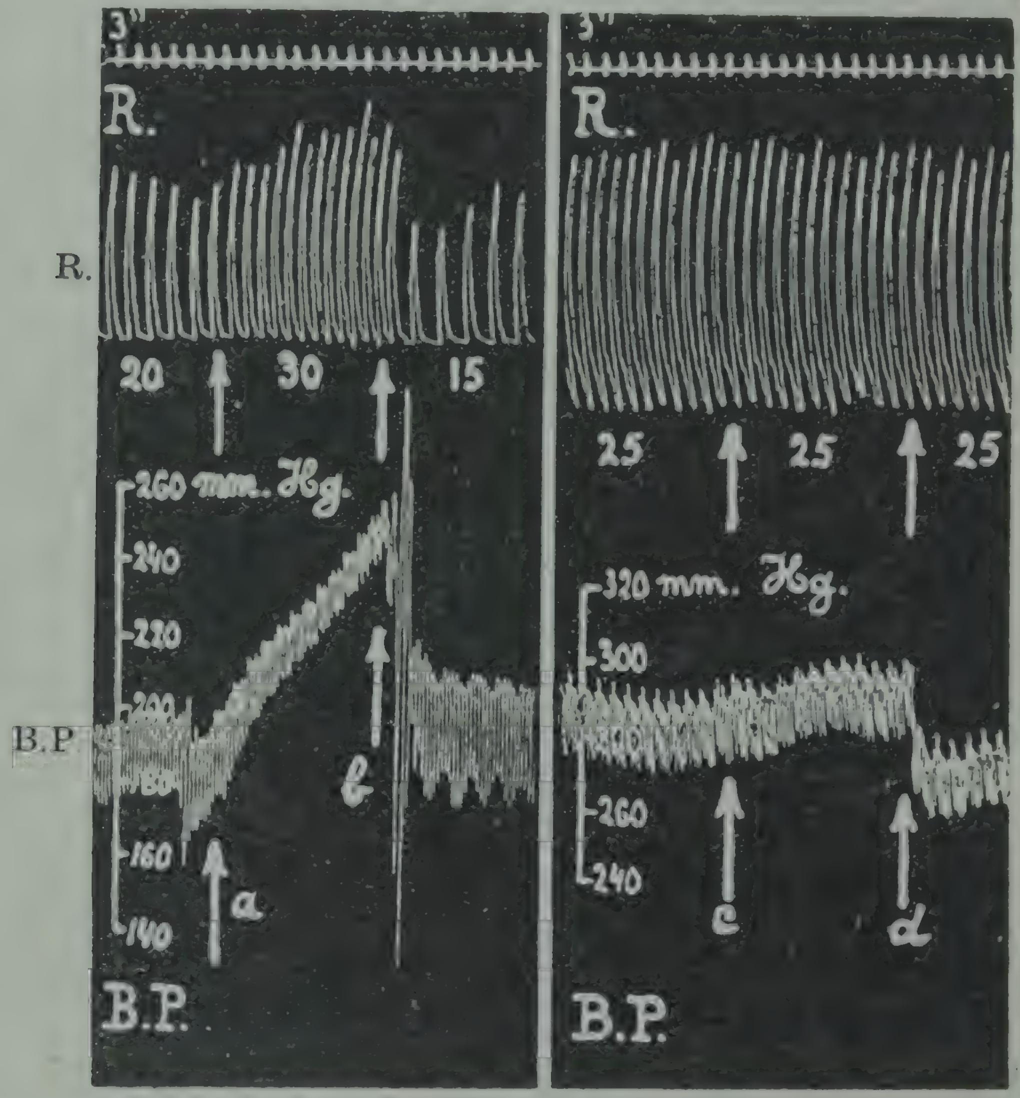
---

* CAMPBELL, DOUGLAS and HOBSON. *J. Physiol.*, 1914, **48**, 303.

† NIELSEN. *Skand. Arch. Physiol.*, 1936, **74**, *Suppl.* 10, 83.

This view would bring under one head all the chemical factors which we know to act upon respiration, namely, tension of carbon dioxide, presence of acids in the blood and considerable diminution of oxygen supply. In a more recent attempt, Gesell * has incorporated the hypotheses of humoral transmission of excitation. According to this view a rise of intracellular *c*H not only excites the cells of the respiratory centre, but also delays the destruction of acetyl choline by which synaptic transmission is effected. Acetyl choline injected intra-arterially causes a normal hyperpnœa, potentiated by eserine, and opposed by atropine, in dogs deprived of chemo-receptors; it also augments the effects of various respiratory reflexes.

Influence of Blood Pressure Changes. Carotid Sinus. It has long been known that the activity of the respiratory centre is modified by alterations in the general arterial pressure. Thus, the occlusion of both carotid arteries leads to a rise of arterial pressure, and to stimulation of respiration. It has been shown by Heymans and Bouckaert † that this effect is due to the reduction of pressure in the carotid sinus, since it disappears if the sinuses are denervated (Fig. 519). They also find that alterations of pressure in the carotid sinuses, when these are isolated from the general circulation and separately perfused, lead to important changes in the respiratory rate (*v.* Fig. 517). A rise of the



A B

FIG. 519. Effect on respiration (R) of alterations of blood pressure of dog produced by occluding both carotid arteries at *a*, in Fig. A, and at *c* in Fig. B. The arteries were released again at *b* and *d*. In Fig. B the carotid sinuses have been denervated, which has caused a rise of respiratory rate by removal of tonic impulses from the sinuses, but change of pressure in the sinuses now produces no further effect. (HEYMANS and BOUCKAERT. *J. Physiol.*, 1930, **69**, 254.)

---

* GESELL *et al.* *Amer. J. Physiol.*, 1942, **136**, 604; 1943, **138**, 776; **139**, 371.

† HEYMANS and BOUCKAERT. *J. Physiol.*, 1930, **69**, 254.

sinus pressure leads to inhibition of respiration, and a fall to acceleration. The afferent nerve for this reflex is the branch of the glossopharyngeal which supplies the sinuses, but the aortic nerves also collaborate with it, as in the regulation of the blood pressure and heart rate, and produce similar effects.

The nerves of the carotid sinuses, and the aortic nerves, must therefore be included with the vagi as exerting a constant effect on the respiratory rhythm *; the effect produced is, however, in the opposite direction. Yet after section of the carotid sinus nerves the respiration is usually depressed. This is because the tonic accelerating impulses arising from stimulation of chemical receptors in the carotid bodies are also cut off.† If, as in Fig. 519, these are not in action, then section of the carotid sinus nerve, by removing inhibitory impulses due to pressure receptors, causes acceleration. We cannot suppose that these reflexes from the carotid sinuses are the only means by which changes in general arterial pressure may affect the respiratory centre, for unquestionably any change which leads to an alteration in the blood supply to the brain will also affect it.‡

Reflex effects on respiration may also originate from the pulmonary vascular bed. Fibres analogous to those of the aortic nerve arise in the pulmonary arteries and veins, and their stimulus by stretching may cause either depression or increase of respiration. The hyperpnœa due to pulmonary embolism is probably an instance of the latter effect.

Apnœa. If artificial respiration be maintained in an animal, or if, in man, the respiratory movements be forced, so as to produce a greater ventilation than is effected by the normal respiratory movements, a standstill of respiration is ultimately brought about. This condition is called apnœa. The standstill cannot be due to over-oxygenation of the blood, because this is normally practically saturated with oxygen, and further, because a short apnœa can even be produced by excessive artificial ventilation with nitrogen.§ It is due to the fact that forced ventilation of the lungs, with any gas mixture free from CO_2, will reduce the carbon dioxide tension in the blood circulating round the respiratory centre.‖ A respiratory pause will thus ensue, and will last until the increasing accumulation of carbon dioxide in the blood raises its tension to the normal height at which the respiratory centre is "set," so to speak, to respond. If the carbon dioxide content of inspired air be increased to about 4·5 per cent., it is impossible to produce an apnœic pause, however rapidly the respiratory movements be carried out. It would seem, therefore, that ordinary apnœa is entirely due to deficiency of carbon dioxide tension in the blood ("hypocarbia" or "acapnia").

It should be noted that over-ventilation and apnœa cannot be caused by artificial respiration by Schäfer's method in the conscious human subject. Indeed, if the subject be brought into an apnœic state by voluntary over-breathing just before applying artificial respiration, it can easily be shown that the same procedure as will ordinarily produce a normal amount of ventilation now induces but little until the state of apnœa has passed off.¶ The reason for this is that the Hering-Breuer reflex is now abnormally sensitive, so that the smallest collapse of the lungs is sufficient to check any further expiratory movement and *vice versâ*, so that the movements of artificial respiration are reflexly resisted.

Stoppage of respiration is sometimes caused reflexly. A familiar example of this is the "*deglutition apnœa*" seen in man during swallowing. It lasts for about 1·5 seconds as a rule, but shows individual variations, and is shorter for swallowing of liquids than of solids. Deglutition generally occurs during the phase of expiration, which is afterwards resumed at the point at which it was arrested.* Similarly, when a duck plunges, there is immediate stoppage of respiration, which may last four or five minutes if the animal remains so long under water. The same stoppage may be produced by pouring water on the beak.

---

* Danielopolu, Marcou and Proca. *J. de Phys. et de Path. Gen.*, 1932, **30**, 632.

† Stella. *Arch. int. Pharm.*, 1936, **53**, 289.

‡ Schmidt. *Amer. J. Physiol.*, 1932, **102**, 94, 119.

§ Haldane and Poulton. *J. Physiol.*, 1908, **37**, 390.

‖ Haldane and Priestley. *J. Physiol.*, 1905, **32**, 258; Douglas and Haldane. *J. Physiol.*, 1908, **38**, 404.

¶ Yandell Henderson. *J. Amer. Med. Assoc.*, 1914, **42**, 1133; Liljestrand *et al.* *Skand. Arch. f. Physiol.*, 1913, **29**, 149.

Breath-holding and Other Tests. After a deep expiration, followed by the deepest inspiration, the breath can be held for an average time of about 68 seconds in fit men,† though the actual times range from 30 to 120 seconds. When the " breaking-point " is reached, it is chiefly because of the accumulation of carbon dioxide in the blood rather than because of lack of oxygen, but the differences are partly due to psychological causes. The alveolar air at the breaking-point contains about 7 per cent. CO_2 and 10 per cent. O_2. Previous forced breathing will enable the breath to be held longer, as also will swallowing, or the filling of the lungs with pure oxygen. In the latter case the alveolar CO_2 may be 8 per cent. or more.‡ Conversely, the period is much shortened if the lungs be previously over-ventilated with nitrogen or with oxygen containing 5 per cent. CO_2.

The " endurance " test or " 40 mm. test " is carried out as follows : the subject's nose is clipped, he takes the deepest possible expiration, followed by the deepest possible inspiration ; the mouth is then applied to a tube connected with a mercury manometer, and he raises the mercury to a height of 40 mm. which is maintained as long as possible, the cheeks and lips being supported by the hand so that they take no part in the process. The pulse rate is counted meanwhile. A fit subject should maintain for 52 seconds, and show no great increase in pulse rate (or arterial pressure) until the point of discomfort is reached. Unfit persons maintain for a shorter time, and show a rise of arterial pressure and pulse rate, at an early stage of the test. Another test is the maximum force of expiration following a deep inspiration, as tested with a mercury manometer ; it should be 110 mm. Hg. These tests, with the requirement of a vital capacity of 3,800 c.c. or over, are employed for the selection of recruits for the Royal Air Force.

The value of the " 40 mm." test is that it tests the extent of abdominal pooling ; when this is large, the intra-abdominal pressure forces out blood from the abdominal area, and increases the return to the heart, so that the arterial pressure and pulse rate rise ; later pulse rate may drop reflexly.§

Adaptation of Breathing in Exercise

As an example of a normal adaptation, we may take the changes in respiration which occur as the result of muscular exercise. The cause of these changes is not entirely explained. During mild activity, the blood will reach the carotid region and respiratory centre with a slightly increased CO_2 tension. The carbon dioxide content of the alveolar air will be raised, but the respiratory centre is stimulated both directly and reflexly, and the increased pulmonary ventilation thereby produced prevents further rise in the alveolar carbon dioxide pressure, until a point is reached at which an equilibrium is maintained between the effect of the increased production of carbon dioxide in raising the arterial carbon dioxide tension and that of the increased respiratory activity in lowering it. Under these circumstances it is found that the increased consumption of oxygen in the contracting muscles is compensated ; in fact, there is a linear relationship between oxygen consumption and the volume of air breathed, when the exercise is moderate in amount.

When the exercise is more severe, however, the respiratory movements are usually increased to such an extent that the tension of carbon dioxide in the

---

[*] CLARK. *J. Physiol.*, 1920, **54**, lix.

[†] SCHNEIDER. *Amer. J. Physiol.*, 1930, **94**, 464.

[‡] HILL and FLACK. *J. Physiol.*, 1908. **37**, 77.

[§] FLACK and BURTON. *J. Physiol.*, 1922, **56**, *Proc.* 1 ; M. R. C. *Special Rep.*, 1920, No. **53**.

blood falls *below* normal. Again, the hydrogen-ion concentration of the arterial blood has been claimed as the decisive factor. If enough lactic acid were produced, the respiratory movements might be augmented to such an extent that the CO_2 tension in the blood and alveolar air falls below normal and yet the hydrogen-ion concentration may be raised. Lastly, there may be deoxygenation of the arterial blood, which would cause further respiratory increase by action on the carotid body.

The fact remains, however, that hyperpnœa often occurs without demonstrable alteration of blood *p*H., O_2 or CO_2 tension. Even *passive* movements of the limbs can cause it, and when this is so, the effect, which can be abolished by section of the spinal cord, appears to be a reflex originating mainly in the joints concerned,* and perhaps in part in the lungs, great veins, and auricles, in consequence of circulatory changes there. In active movements of the limbs, pressor reflexes from chemo-receptors in the limbs may be set up,† and no doubt irradiation of excitation from cortico-spinal paths into the reticular formation of the medulla is a contributory factor. Rise of blood temperature, acting on carotid bodies and on nerve centres may also play a part.

Liberation of adrenaline is another factor that may augment ventilation. In all probability the hyperventilation of exercise cannot be explained as due to any single factor.

Asphyxia (*v.* p. 754). The first phase in the phenomena of *asphyxia* is conditioned simply by the changes in the carbon dioxide tension. A little later, the gradual exhaustion of oxygen in the blood begins to make itself felt, and the reflexes from the aorta and carotid body, due to this cause, further augment the breathing. During the second stage there is a discharge of energy throughout the whole nervous system, beginning in the bulbar centres and causing a great rise of blood pressure, with slowing of the heart, and extending thence to all the spinal centres with the production of muscular spasms. At this stage, too, there is a discharge of impulses causing contraction of the pupil. The phenomena of the third stage are due to the influence of oxygen lack upon the nerve centres, accompanied or preceded by exhaustion and dilatation of the heart, the circulation failing before the excitation of the lower centres has entirely come to an end. In this third stage it is impossible, even by the strongest stimuli, to evoke any reflex.

Dyspnœa. We are not generally aware of an increase in the volume of the pulmonary ventilation until this has been about doubled. With further increase the consciousness of respiration steadily obtrudes itself, until by the time the ventilation has been increased to about four or five times the normal resting value, there is real distress, and we say that dyspnœa has set in. Dyspnœa is seen during strenuous exertion, in conditions of oxygen lack, and in certain diseased states.

The dyspnœa of exercise is associated with deficient oxygenation and increased carbon dioxide content of the blood, and these in turn may be said to be due to an improper adjustment between circulation and respiration. Usually it is the heart which is at fault, and which is the limiting factor in athletic performances of endurance ; the heart is unable to pump the blood through the lungs quickly enough, so that it becomes increasingly venous and thus excites the respiration more and more.

If vigorous exercise is continued, there comes a time when the discomfort disappears, and the subject gets what is called "second wind." It is found that at this point there is a sudden fall in the carbon dioxide content of the alveolar air ; during the dyspnœic period this may have been as high as 7·5 per cent., which often falls after second wind to 6·0 per cent. At the same time there is a fall in the volume of air breathed, a rise of temperature, and onset of sweating.‡

---

* Comroe and Schmidt. *Amer. J. Physiol.*, 1943, **138**, 536.
† Alam and Smirk. *J. Physiol.*, 1937, **89**, 372.
‡ Cook and Pembrey. *J. Physiol.*, 1913, **45**, 429.

The phenomenon has not received a full explanation; it is probable that it represents the attainment of a steady state, in which the circulation rate, the chemical composition of the blood, the psychical condition of the individual, the state of the respiratory centre, of heat regulation, etc., reach new equilibria.

The dyspnœa in normal subjects associated with oxygen lack will be specially considered later.

Pathological causes of dyspnœa can only briefly be mentioned. Apart from those, such as pneumothorax, obstruction to the air-passages, etc., which are due to purely mechanical causes opposing the entry of air, the chief types are those seen in renal disease, cardiac disorders, diabetes and anæmia. In renal disease* and in diabetes, there is an excess of fixed acids in the blood, raising its hydrogen-ion concentration. In cardiac cases there is not only a reduced circulation rate, but usually some secondary alterations in the lungs themselves, for both of which reasons there is difficulty in keeping the blood properly arterialised. In anæmia, the number of red corpuscles in the blood is insufficient to carry the proper amount of oxygen to the tissues, so that the condition is virtually the same as in other states of oxygen lack. All these conditions, however, are actually more complicated than would appear from what has been said above, and often the cause of the dyspnœa remains obscure.†

Over-ventilation. (Forced Breathing.) There is a large amount of carbon dioxide, free or fixed, stored in the blood and tissues, and we may compute the whole body of an adult man to contain some 15 or 20 litres of carbon dioxide. The oxygen store of the body is held almost entirely by the hæmoglobin, and does not exceed 1 litre. Now, forced (*i.e.* voluntarily quick and deep) breathing moves large amounts of air; in fact the maximal ventilation so obtained can exceed that of exercise. This hyperventilation very rapidly washes out carbon dioxide from the large store, but does not at all increase the oxygen store. By continuing forced respiratory movements for a minute or two, the carbon dioxide tension, both in the alveoli and in the blood, may be brought down very considerably.

In consequence of the lowering of carbon dioxide pressure in the blood, this becomes more alkaline, and various secondary phenomena ensue. Thus, alkaline urine is excreted, and this always contains bicarbonate; it also contains less ammonia than previously and may contain aceto-acetic acid.‡ The bicarbonate of the blood plasma is lowered,§ partly by reason of the ionic interchange with the corpuscles, partly by excretion of base. The blood lactate is increased, because glycolysis in the blood is accelerated. The calcium ion concentration of the blood is lowered, and in consequence symptoms of tetany may appear. The rise in pH of the blood, together with the fall of CO_2 pressure, affects the vasomotor centre, causing a fall of arterial pressure,‖ and also depresses the respiratory centre. As a result, there is a prolonged period of

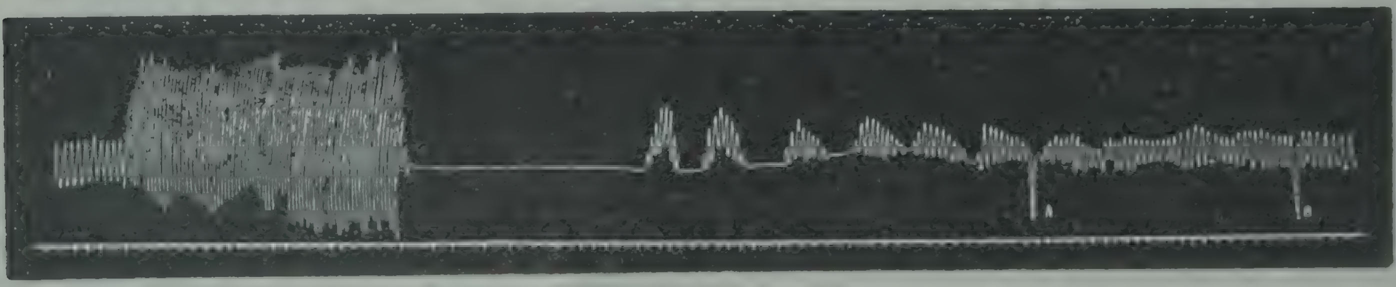


Fig. 520. Forced Breathing of Air for two minutes, followed by Apnœa for two minutes, and Periodic (" Cheyne-Stokes ") Breathing for about five minutes. (Douglas and Haldane.)

At A, sample of alveolar air contained O_2, 11·44 per cent.; CO_2 5·58 per cent. Second sample at B, O_2, 13·55 per cent.; CO_2 5·57 per cent.

---

* Henderson *et al.* *J. Biol. Chem.*, 1927, **75,** 305.

† Campbell, Hunt and Poulton. *J. Path. and Bact.*, 1923, **26,** 234; Henderson. "Blood," 1928, p. 265 (Yale Univ. Press); Dautrebande. "Les échanges respiratoires." 1930, pp. 105–282 (Paris: Press Universataires de France).

‡ Davies, Haldane and Kennaway. *J. Physiol.*, 1920, **54,** 32.

§ Henderson and Haggard. *J. Biol. Chem.*, 1918, **33,** 333; Dale and Evans. *J. Physiol.*, 1922, **56,** 125.

‖ Henderson and Haggard. *J. Biol. Chem.*, 1918, **33,** 333; Dale and Evans. *J. Physiol.*, 1922, **56,** 125.

apnœa. During this period of cessation of respirations, however, the oxygen is being used up, and the tension of this gas in the blood may fall to such an extent that the respiratory centre is reflexly excited by lack of oxygen before the carbon dioxide tension in the alveoli has risen to its normal value. As a result of the excitation by oxygen lack, a few breaths are taken, the carbon dioxide tension is further lowered, and the stimulation due to the oxygen lack disappears. There is thus again a cessation of respiration (Fig. 520).

"Cheyne-Stokes" Breathing. These periods of cessation alternate with periods of respiration, so that we get a condition of periodic breathing, which is spoken of as Cheyne-Stokes respiration.* During the period of apnœa resulting from forced breathing, the great diminution of oxygen tension in the alveoli is shown by the fact that the subject of the experiment becomes blue, and may indeed lose consciousness. There are, at the same time, rhythmic changes in the blood pressure, which rises towards the ends of the periods of apnœa, falling during the periods of respiration. The first respiration after excessive forced breathing is due to oxygen lack. The period of apnœa may therefore be considerably prolonged, if the onset of oxygen lack be postponed by increasing the tension of this gas in the alveoli at the commencement of the apnœic period. By forcibly breathing for a period of two minutes in an atmosphere of oxygen, men have succeeded in holding their breath for as long a period as eight minutes.†

"Cheyne-Stokes" breathing is almost invariably observed as one of the effects of exposure to high altitudes, and is then especially marked during sleep. It is often present when the activity of the respiratory centre is depressed, as in cases of uræmia or pernicious anæmia. Under these circumstances, it may be temporarily removed by administering either oxygen or carbon dioxide (in small percentage) to the patient. The oxygen removes the anoxic stimulus to the carotid body; the carbon dioxide prevents the hypocarbic apnœa.

ANOXIA

The name anoxia is given to the condition of oxygen lack of the tissues. Lack of oxygen in the blood is also called *anoxæmia*. Four types of anoxia are recognised, viz. *anoxic, anæmic, stagnant* and *histotoxic*.

Anoxic Anoxia is anoxæmia, *i.e.* a lowered partial pressure of oxygen in the blood, with consequent diminution in the degree of oxygen saturation of the arterial blood. It is produced by any condition which prevents the proper oxygenation of the blood in the lungs, *e.g.* by reduced oxygen pressure in the air breathed, by diseases in which the gaseous interchange in the lungs is impaired, and in congenital heart disease due to a patent foramen ovale.

Anæmic Anoxia. In this form the oxygen tension of the blood is unchanged, but the quantity of oxygen which can be made available from the blood is diminished. It is seen in hæmorrhage, in anæmia, and in states in which the hæmoglobin is converted into some form which will not carry oxygen, *e.g.* into nitric oxide hæmoglobin, as in poisoning by nitrous fumes, or into methæmoglobin, as in poisoning by aniline or chlorates, or into carboxyhæmoglobin in carbon monoxide poisoning.

Stagnant Anoxia is seen in conditions in which the circulation through a part or over the whole body is slowed. Thus it is seen in surgical shock, circulatory failure, and venous obstruction.

Histotoxic Anoxia results when, the oxygen pressure and content of the blood being unaltered, or even raised, the tissues, though in full receipt of blood, fail to utilise the oxygen from it. In this category may be placed the poisoning by cyanides, which paralyse the respiratory enzymes of the tissues, and probably also the toxic effects of very high oxygen pressures.

---

* Cheyne. *Dublin Hosp. Rep.*, 1816, **2**, 217; Douglas and Haldane. *J. Physiol.*, 1909, **38**, 401.

† Vernon. *J. Physiol.*, 1909, **38**, Proc. xviii.

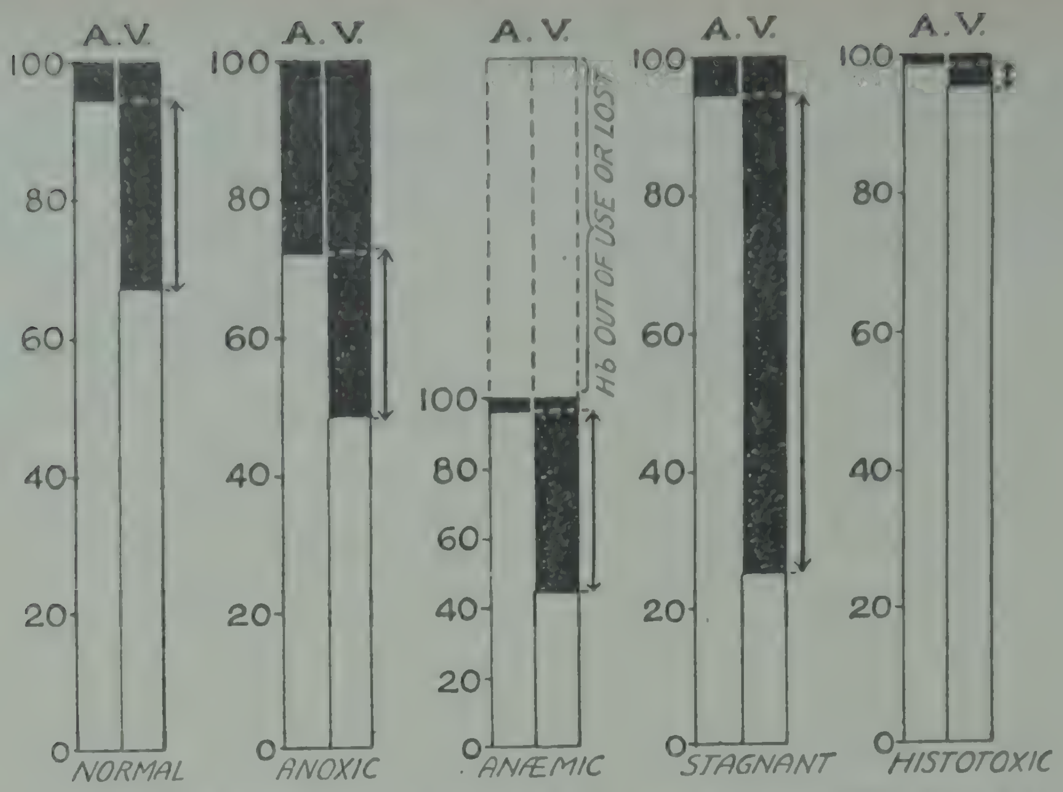


Fig. 521. Oxygen saturations of arterial (A) and venous (V) blood in the four types of anoxia. The black portion of each column represents the percentage of reduced hæmoglobin and the white portions the percentage of oxyhæmoglobin in the blood. All the bloods had the same initial oxygen capacity when the subject was normal. In anæmic anoxia some of this, shown by dotted lines, has been lost, *e.g.* by hæmorrhage or by conversion to carboxyhæmoglobin. The arrows show the coefficient of oxygen utilisation, *i.e.* the volume of oxygen removed from one unit of blood by the tissues.

The oxygen saturations of the arterial and venous bloods in the four types of anoxia are illustrated in Fig. 521.

CYANOSIS

Cyanosis is the name given to a bluish colour of the skin, especially of the lips, ears, hands and feet, and of the nails. It is due to the presence of reduced hæmoglobin in the blood, and so is met with in conditions in which this occurs, whether from inadequate oxygenation of the arterial blood, or from the utilisation of an undue proportion of the oxygen of the blood during its transit through the capillary circulation, as in sluggish local or general circulation.

The degree of cyanosis depends, not on the relative amount of reduced hæmoglobin in the capillary blood, but on its *absolute amount.* To cause cyanosis there must be at least 5 g. of reduced hæmoglobin per 100 c.c. of capillary blood. Since blood normally contains about 15 g. Hb per 100 c.c., this means that less than 66 per cent. oxygen saturation of the blood in the capillaries will cause cyanosis. But an anæmic person with half the normal Hb content would have to be only 33 per cent. saturated with oxygen before cyanosis was evident.

It is convenient in this connection to consider the *degree of unsaturation* of the blood. Normal arterial blood is about 5 per cent. unsaturated, which for an oxygen capacity of 20 vols. per cent. is 1 c.c. oxygen deficit or unsaturation per 100 c.c. blood. The mixed venous blood is about 6 c.c. per 100 c.c. unsaturated, *i.e.* the tissues have taken out 5 c.c. oxygen per 100 c.c. blood. The mean capillary unsaturation may be taken as midway between the arterial and venous levels, *i.e.* as 3·5 c.c. per 100 c.c. Since 1 c.c. oxygen will fully saturate 0·75 g. Hb, the capillary blood will therefore contain $3 \cdot 5 \times 0 \cdot 75 = 2 \cdot 5$ g. Hb in reduced form, *i.e.* there will normally be no cyanosis.

Suppose now that we consider the blood in a subject in a state of anoxic anoxia, due, suppose, to breathing an atmosphere in which the oxygen had been reduced. The blood leaving the lungs contains, say, only 15 vols. oxygen per 100 c.c.; the tissues abstract the usual quota of 5 vols. per cent., so that the venous blood contains 10 vols. per 100 c.c. The arterial and venous unsaturations are respectively 5 and 10 c.c. and the capillary unsaturation again the mean of these, viz. 7·5 vols. per cent., = 7·5 × 0·75 = 5·6 g. Hb. There would be cyanosis.

Again, suppose the arterial blood to be normal, *i.e.* 1 vol. per cent. unsaturated, but owing to slow circulation the venous blood was 15 vols. per cent. unsaturated. The capillary unsaturation would be 8 vols. per cent. = 8 × 0·75 = 6 g. Hb, and again there would be cyanosis.

An anæmic subject with an oxygen capacity of half the normal, viz. 10 vols. per cent., if exposed to the anoxic conditions would have 0·5 vol. per cent. unsaturation in his arterial blood, and 9·5 — 5·0 = 4·5 vols. O_2 or 10 — 4·5 = 5·5 vols. unsaturation per cent. in his venous blood. His capillary unsaturation would thus be the mean of 0·5 and 5·5, *i.e.* 3·0 c.c. = 3 × 0·75 = 2·25 g. Hb, so there would *not* be any cyanosis.

In the reverse condition of polycythæmia, or in states of "hæmoconcentration," in which the number of red cells is abnormally increased, the opposite would hold good, and cyanosis would ensue on relatively slight anoxæmia.

For Reference

Comroe (1944). The Hyperpnœa of Muscular Exercise. *Physiol. Rev.*, **24**, 319.
Gesell (1925). The Chemical Regulation of Respiration. *Physiol. Rev.*, **5**, 551.
Pitts (1946). Organisation of the Respiratory Centre. *Physiol. Rev.*, **26**, 609.
Pi-Suñer (1947). The Regulation of the Respiratory Movements by Peripheral Chemoreceptors. *Physiol. Rev.*, **27**, 1.

THE EFFECTS ON RESPIRATION OF CHANGES IN THE AIR BREATHED

We have already seen that a moderate increase in the carbon dioxide content in the air breathed (*e.g.* up to 4 per cent.) causes a proportional increase in the ventilation of the lungs and that this is due to the increase of CO_2 and not to the slight diminution in oxygen content. High percentages of CO_2 cause abnormal effects leading to narcosis and death.

Changes in Oxygen Tension. Oxygen usually exercises little, if any, effects on the respiratory ventilation. At the normal atmospheric pressure, the partial pressure of oxygen in the alveoli, when air is breathed, is amply sufficient to nearly saturate the hæmoglobin passing through the lungs. Hence little reduction in total ventilation will be produced by increasing the partial pressure of oxygen in the alveoli, even when pure oxygen is inhaled. But prolonged breathing of pure oxygen may cause lung damage.

This statement is true only for the healthy individual. If, from failure of the heart and circulation, from diminished oxygen partial pressure, or from severe loss of blood, the oxygenation of the blood is insufficient, amelioration of the symptoms may be produced by inhalation of pure oxygen. Especially is this noticeable where there is failure of the heart. In these cases, the heart, already affected, is unable to keep up an adequate circulation and to supply itself with sufficient oxygen. A vicious circle is thus established, in which the heart tends to get steadily worse. By administration of oxygen, an adequate supply of this gas to the heart muscle is assured; the heart beat therefore becomes more effective and the whole circulation is improved.

Breathing of oxygen at high pressures is an altogether different matter. If a warm-blooded animal be placed in a chamber containing pure oxygen at four atmospheres, or air at twenty atmospheres, it dies as rapidly as if it were placed in nitrogen.* If removed while living, recovery of the animal may be rapid, though after-effects on the nervous system may ensue. This fact assumed importance in naval diving practice † after the introduction of closed-circuit oxygen-breathing apparatus. The pressure rises, as the diver descends, at the rate of nearly 1/5 atm. per fathom, so that, *e.g.* at 90 ft. the pressure is raised by 2·73 atm. (= 3·73 atm. absolute). The time of onset, and the nature and severity of symptoms varies from subject to subject, and from time to time; but almost every subject will, at 90 ft. depth, be affected within the hour, and some in a few minutes. Some risk may be expected in sea-diving with oxygen at any depth below 25 ft. of sea water (= 1·75 atm. abs.). The symptoms include psychological changes, twitching of lip, facial and other muscles, and finally epileptiform convulsions and loss of consciousness.

---

* Paul Bert. *La Pression Barométrique.* 1878 (Paris); translation, 1943, *Barometric Pressure* (Columbus: College Book Co.); Lorrain Smith. *J. Physiol.*, 1899, **24**, 19.

† Donald. *B.M.J.*, 1947, *i*, 667, 712.

Two explanations of the phenomena are offered, and probably both are partly correct:[*] at the high pressure enough oxygen can be taken up by the animal's *plasma* to supply its wants, so that the oxyhæmoglobin does not become reduced. But then, the more acid oxyhæmoglobin retains the available base of the blood and so prevents this base from combining with CO_2 from the tissues. Hence there is a rise of CO_2 pressure within the tissues, which ultimately causes death. As pointed out by Campbell, it is the rise of CO_2 pressure in the tissues which causes the convulsions that precede death under high oxygen pressures. There is also some evidence that high oxygen pressures poison respiratory enzymes, *e.g.* dehydrogenase or, more particularly pyruvic oxidase, by attacking the —SH group of the enzyme.

We have seen that the oxygen pressure in the air may be lowered to about 13 per cent. without pronounced discomfort, though there may be blueness of the lips and nails, owing to deficient oxygenation of the blood. With further reduction, hyperpnœa occurs, owing to excitation of chemo-receptors in the carotid glomus, and, if the administration of low percentages of oxygen, *e.g.* about 10 to 12 per cent. of an atmosphere, be continued for some time, the subject of the experiment may suffer considerable discomfort. The symptoms of oxygen lack are severe headache, confusion of thought, vomiting or nausea and a feeling of discomfort in the precordial region.

Respiration at reduced Atmospheric Pressure. Many experiments have been made, on animals and on men, submitted to a lowered atmospheric pressure in chambers specially built for the purpose.[†] The limit to which the pressure may be reduced varies in different individuals, the variations being determined by the type of respiratory movement of the individual in question, since on the depth of respiration depends the relation between the tensions of oxygen in the alveoli and in the inspired air. The lowest limit at which human life is possible corresponds to an oxygen tension in the alveoli of about 25 mm. Hg. The physiology of respiration at reduced atmospheric pressures is particularly interesting in connection with exploits in mountaineering or aviation; the chief difference between the two being that in the former great physical exertion must be coupled with the drawbacks of reduced air pressure and adverse climate, while in the latter there is the drawback that the reduction of pressure is rapid and the onset of symptoms correspondingly sudden. A good deal of investigation has been carried out for the purpose of studying the phenomena involved in the response and acclimatisation to lowered air pressure, during both aircraft[‡] and mountain ascents.[§]

The phenomena of oxygen lack above described are similar to those which are regarded as characteristic of "mountain sickness." The Table on p. 775 shows the diminution in the atmospheric pressure at varying heights above the level of the sea.

At a height of 5,000 metres (16,400 feet) the pressure of the air is reduced to little over half an atmosphere, and the oxygen tension is therefore only about 11 per cent. of an atmosphere. Moreover a greater *volume* of the alveolar air

[*] Campbell. *J. Physiol.*, 1925, **60,** 20; 1929, **68,** *vii*; Bean. *J. Physiol.*, 1931, **72,** 27; Behnke *et al.* *Amer. J. Physiol.*, 1934, **107,** 13.

[†] Boycott and Haldane. *J. Physiol.*, 1908, **37,** 355; Haldane, Kellas and Kennaway. *J. Physiol.*, 1919, **53,** 181; Schneider, Truesdell and Clarke. *Amer. J. Physiol.*, 1924, **70,** 283.

[‡] Schneider and Clarke. *Amer. J. Physiol.*, 1926, **76,** 354.

[§] Douglas, Haldane, Henderson and Schneider. *Phil. Trans.*, 1913, **B203,** 185 (Pike's Peak); Barcroft *et al.* *Phil. Trans.*, 1922, **B211,** 351 (Andes); Hingston. *Geogr. J.*, 1925, **65,** 4 (Everest Exped.).

Height above sea level (in metres)	(in feet)	Barometer mm. Hg.	Pressure in per cent. of an atmosphere
0	0	760	100
1,000	3,281	670	88
2,000	6,562	593	78
3,000	9,842	524	69
4,000	13,123	463	61
5,000	16,404	410	54
6,000	19,685	357	47
7,000	22,966	320	42
8,000	26,246	290	38
9,000	29,523	260	34

must consist of carbon dioxide since the *tension* of this gas tends to be kept nearly constant (cp. Fig. 508).

Because of this dilution with CO_2, and also with water vapour, at reduced atmospheric pressures, the effects of a lowering of oxygen content of air at normal pressure are not the same as those of a proportional lowering of air pressures, *e.g.* 10·5 per cent. oxygen at 760 mm. has not the same effect as 21 per cent. at 380 mm., because if p $CO_2 = 40$ mm., and p $H_2O = 47$ mm., then the oxygen pressures in the two cases are :—

$$(760 - 40 - 47) \times 0{\cdot}105 = 71{\cdot}7 \text{ mm. Hg.}$$
$$(380 - 40 - 47) \times 0{\cdot}21 = 61{\cdot}5 \text{ mm. Hg.}$$

This means to say that, in laboratory experiments, a reduction of O_2 content of the air to one-half of the normal corresponds to an altitude of only 17,000 ft. and not to 18,000 ft.*

The much greater diminution in the partial pressure of oxygen than in that of CO_2 in the alveolar air is shown by the following average results obtained during fourteen aeroplane flights, with five subjects :—

ALVEOLAR AIR DURING AEROPLANE FLIGHTS. (SCHNEIDER and CLARKE. *Amer. J. Physiol.*, 1926, **76**, 354.)

Altitude, feet	0	4,000	8,000	12,000	15,000
Alveolar oxygen pressure mm. Hg .	100·5	90·5	67·2	58·1	49·8
Alveolar CO_2 pressure mm. Hg . .	40·9	35·3	35·7	32·5	29·2

The carbon dioxide pressure, however, does fall somewhat, and this is due to the increased pulmonary ventilation, which shows a steady increase as the reduction of pressure progresses. If the low pressure is maintained, the breathing continues steadily to rise to a maximum ventilation some hours or days later. In the 1938 Mt. Everest expedition the alveolar partial pressure at 22,800 feet was 16·5 mm. CO_2 and 35·7 mm. O_2 after acclimatisation ; at 28,000 feet (reached) they would probably be 16 mm. and 25 mm. respectively. The increased breathing is a consequence of the reduced oxygen pressure and provides an increased lung surface, and also an increased excretion of carbon dioxide, and so leads at first to an increase in the blood alkalinity and to the excretion of an alkaline urine. If the volume of air breathed be reduced by calculation to normal pressures, it is found that less air is breathed than at sea level.

Deep breathing affords a means of immediate adaptation, and is also operative in the adaptation which comes on gradually during residence at a high altitude. Part of this acclimatisation appears to depend upon the excretion, by the kidneys, of bicarbonate from the blood, thus compensating for the low content of free CO_2. When this has been brought about, the pH of the blood returns practically to normal.

---

* HORVATH, DILL and CORWIN. *Amer. J. Physiol.*, 1943, **138**, 659.

Another important means of rapid adaptation is by means of the circulation. This is noticeable even in persons sitting quietly in a chamber in which they can be subjected to gradually lowered pressures, and no doubt holds good for aviators. It is evident that a deficient oxygenation of the blood may be compensated by increasing the rapidity of the circulation ; and this happens. Both the stroke volume of the heart and the pulse rate are increased in the early stages of oxygen deprivation. The following Table shows the changes in the pulse rate caused by exposure to varying pressures in a gas chamber :—

PULSE IN CHAMBER AT REDUCED PRESSURES

Pressure mm. Hg.	Pulse p. min.
720	64
650	72
424	84

This quickening of the pulse is to be observed also in the trained mountain soldier, and in individuals in whom there is no lowering of the alveolar carbon dioxide tension, so that apparently, in such cases, the whole adaptation to altered conditions is by means of the circulation. In cases where adaptation fails, it is in the circulation that the failure is most marked, so that the symptoms of severe mountain sickness closely resemble those produced by heart failure. But this mechanism, again, is only immediate and does not explain acclimatisation to life in mountain districts. Barcroft, in the Peruvian Andes (4,400 metres) found that at rest the pulse was normal and that there was no increase in the circulation rate.

As a result of extreme oxygen starvation, there may be an inadequate supply of this gas to the heart, which may tend to fail, especially on making muscular movements. But at a much earlier stage, the oxygen starvation of the brain produces failure of judgment, inability to carry out or to co-ordinate muscular movements properly, and then loss of consciousness. The disturbance of the central nervous system is shown by the almost invariable occurrence of Cheyne-Stokes breathing at great heights. The symptoms of oxygen lack, as a rule, do not become conspicuous until just before collapse, so that, although there is an oxygen starvation of the body, there must be some compensation to obtain a sufficiency of this gas for the lowered requirements of the body. As the subject of oxygen lack is unaware of these effects, one of which takes the form of extreme stubbornness, the insidious danger of the condition is of great importance in aviation. In rapid lowering of oxygen, or in rapid ascents consciousness is lost at 4·2—5·2 per cent. O_2, which corresponds to an altitude of 28,000 to 31,000 feet, but is soon regained on rise of pressure.* Breathing of oxygen-enriched air is essential for pilots at altitudes over 15,000 feet, if effects of oxygen lack are to be avoided. At over 43,000 feet even this does not suffice, as the barometer is too low to give the requisite oxygen pressure ; hence the pressure must be raised by the use of either a sealed flying suit filled with oxygen under slight pressure or a sealed cockpit filled with oxygen—also under slight pressure. With such arrangements, oxygen lack ceases to be a limiting factor in high flying, and heights of about 60,000 feet have been officially recorded.† In jet-propelled aircraft, thus equipped, it would seem that there would be no limit, mechanical or physiological, to the altitude attainable. An interesting point concerning this is the action to be taken by aircrews baling out at, say, 30,000 feet. Should a delayed drop be made, to try to clear the more rare levels before consciousness is lost, or should they open out as soon as free, and lose and regain consciousness during descent ? It seems to be a matter of taste, as well as of courage. Bale-out oxygen sets, with 10 minutes' supply provide the practical answer.

If a subject is able to withstand the immediate effects of exposure to a rarefied atmosphere, processes of adaptation come into play which finally

* HORVATH, DILL and CORWIN. 1943, *loc. cit.*

† Group-Capt. J. CUNNINGHAM, 59,492 feet (25 March, 1948).

fit him for discharging his functions normally, even at the high altitude. That the adaptation is effective, is shown by the fact that most individuals, if they remain at a height, gradually recover from the mountain sickness and may finally be able to carry out muscular movements with almost as great precision and force as they could previously on the plains. The main factor in rapid adaptation to lowered oxygen pressure is not an increase in the tissue O_2-tension, but an habituation of the heart and nervous tissues to the lowered O_2-tension.* As an immediate effect of change to a region of low atmospheric pressure, there is an apparent increase in the blood corpuscles, due to a concentration of the blood and a diminution of its plasma. Simultaneously, however, the blood-forming organs respond to the reduction in the oxygen saturation of the blood, so that, after a stay of four or five weeks' duration at a height, both corpuscles and hæmoglobin are considerably increased in total amount. The following Table shows the average number of red corpuscles contained in one cubic millimetre of blood from inhabitants of regions at varying altitudes :—

	Height above sea level, in metres	Red corpuscles
Christiania . .	0	4,970,000
Zurich . .	412	5,752,000
Davos . .	1,560	6,551,000
Arosa . .	1,800	7,000,000
Cordilleras .	4,392	8,000,000

As a result of the increase in the concentration of red cells, there would, in the later stages of acclimatisation, be an increase in the tissue O_2-tensions.†

At the same time, owing to a rise in the excitability of the respiratory centre, the total ventilation is increased, so that the alveolar oxygen tension is raised 10 or 12 per cent. above that which it would otherwise be.

There is, of course, a limit to the power of adaptation, a limit which varies in different individuals. Thus, for some men, it is impossible to stay any length of time in the high settlements in the Andes, while others, after two or three weeks' discomfort, become perfectly inured to their new conditions. It seems doubtful, however, whether any of the present race of men could become adapted to permanent residence at a height over 6,000 metres, though for a certain length of time, by bringing into play the reserve mechanisms already described, they may raise themselves to a height considerably above 7,000 metres. The highest summits in the Himalayas (*e.g.* Mount Everest, 8,865 metres), have a height approaching that attained in 1875 by Tissandier with two companions in his famous balloon ascent, namely, 8,600 metres. In this ascent, although oxygen inhalation was used (somewhat ineffectively), two of the party succumbed. The expeditions to Mount Everest showed, however, that when acclimatisation was gradual, men could live at about the same altitude, *e.g.* Mallory's party reached 8,200 metres (26,900 feet) without oxygen inhalation. But at such heights, the slightest exertion, *e.g.* changing from lying to sitting, caused extreme breathlessness. Oxygen often gives relief under these conditions, though its administration presents difficulties.‡

For Reference

BEAN (1945). Effects of Oxygen at increased Pressure. *Physiol. Rev.*, **25**, 1.
BEHNKE (1941). Physiologic Studies Pertaining to Aviation Medicine and Deep Sea Diving. (*Harvey Lects.*, 1941-42: pp 198.)
DILL (1938). Life, Heat and Altitude. (London: Oxford Univ. Press.)
GEMMILL (1943). Physiology in Aviation. (Baillière.)
MONGE (1943). Chronic Mountain Sickness. *Physiol. Rev.*, **23**, 166.

---

* CAMPBELL. *J. Physiol.*, 1926, **62**, 211; 1927, **63**, 325.
† CAMPBELL. *J. Physiol.*, 1928, **65**, 256.
‡ LLOYD. *Nature*, 1939, **143**, 961.

Respiration at High Pressures. New phenomena appear when the air pressure is greatly increased, as in deep diving, in caissons, or with those using submarine escape apparatus (*within* a submarine the pressure is normal). When the air pressure is raised to 8·6 atm. (= 250 ft. sea-water) or more, the *nitrogen exerts an intoxicating effect* within a few minutes; added to this is an enhanced effect of CO_2.[*] At normal pressures, 0·8 per cent. CO_2 causes no detectable mental impairment; if, however, the air containing it is compressed to 10 atm. (= 300 ft. sea-water), the partial pressure of CO_2 will be = 8 per cent. at sea-level, and for some reason this is more effective still at high pressure, since men become confused and clumsy, and mostly lose consciousness in a few minutes. The nitrogen of air is normally a purely inert gas. The very inertness of nitrogen may be of danger, when high pressures are reduced, since nitrogen is especially soluble in adipose and nervous tissues.[†] In all, the body normally contains about 1 litre of dissolved nitrogen, one-fifth of which can be washed out in a few minutes by breathing oxygen.[†] If a man be exposed, as in a caisson, to a pressure of three to six atmospheres, the nitrogen dissolved in the body is increased in direct proportion to the pressure. If the pressure be now suddenly released, the nitrogen, which cannot be used up by the tissues, is given off from the tissues in the form of insoluble bubbles. These bubbles, occurring in all the capillaries, obstruct the flow of blood, and therefore, if the evolution of gas is sufficiently large, the animal dies of gas-embolism in heart and nervous system. During the decompression of divers and caisson workers, or in steep high airplane climbing, a similar evolution of gas may occur, causing damage to the cord and paralysis (" caisson disease "; " bends," " divers' palsy ").[‡] In order to prevent this sudden evolution of gas, it is necessary that the decompression should be carried out gradually (best of all in oxygen), so as to give the tissues time to get rid of their excess of nitrogen without the formation of bubbles.[§] Another method is to replace the atmospheric nitrogen by helium, giving the diver a helium-oxygen atmosphere, since helium is less soluble in fat than nitrogen.[||] By this means a depth of 535 feet has been reached (=17·2 atm.).[¶]

OTHER GASES. Argon, helium, hydrogen and methane are, like nitrogen, indifferent gases, and may without harm be used, instead of nitrogen, to dilute the oxygen that we breathe.

Carbon monoxide is rapidly poisonous since it combines with hæmoglobin, forming CO-hæmoglobin. The blood is therefore deprived of its oxygen carrier, and the animal dies of asphyxia. The importance of CO poisoning lies in the fact that coal gas, motor car exhaust, breech gases from guns, etc., contain a high proportion of CO. The blood of town-dwellers usually contains 0·1—0·2 vols. p.c. of CO. Not more than 0·05 p.c. CO in air can be breathed for an hour or two without effects. Breathing air containing 0·1 per cent. of CO causes severe symptoms in thirty to sixty minutes, with per cent. saturations of the Hb with CO of the order of 15 to 20; at 30 to 40 per cent. saturation consciousness is suddenly lost, and death may follow at 50 per cent. saturation or over. In such chronic poisoning with CO, there is a great lowering of the tissue oxygen-tensions.[**] But, although the avidity of CO for hæmoglobin is about 300 times that of oxygen, we can convert the CO-hæmoglobin back into oxyhæmoglobin by increasing the mass influence of the oxygen and by adding CO_2. This may be done by giving the poisoned animal oxygen containing 5 per cent. CO_2 to breathe.

---

[*] Case and Haldane. *J. Hygiene*, 1941, **41**, 225.

[†] Campbell and L. Hill. *J. Physiol.*, 1931, **71**, 309.

[‡] Hill and Macleod. *J. Hygiene*, 1903, **3**, 436; Hill and Greenwood. *Proc. Roy. Soc.*, 1907, **79B**, 21.

[§] L. Hill and Macleod. *J. Physiol.*, 1903, **29**, 382.

[||] End. *Amer. J. Physiol.*, 1937, **120**, 712.

[¶] Petty Officer W. Bollard, R.N. (28 August, 1948).

[**] Campbell. *J. Physiol.*, 1929, **68**, 81.

The curve relating CO-saturation to CO-pressures is similar in form to the oxygen dissociation curve, but, as shown by Douglas, Haldane and Haldane,* the CO-pressures required are very small, only about $\frac{1}{245}$th as great as the oxygen pressures. Thus, we get half saturation of the blood with CO at a CO-pressure of about 0·13 mm. Hg (in presence of 40 mm. CO_2), whereas for oxygen we require a pressure of about 32 mm. (with 40 mm. CO_2). When, however, oxygen and carbon monoxide are both present, the saturating power of carbon monoxide is not 245, but 310, times as great as oxygen. This is possibly due to the fact that the velocity constant for oxygenation of hæmoglobin is about 7·5 times as great as that for reduction, together with the fact that, although the rate of combination of CO with hæmoglobin is only one-tenth as great as for oxygen, the rate of dissociation of COHb is only $\frac{1}{245}$ of that of HbO_2. Hence the stability of COHb is due, not to the speed of its formation, but to the slowness of its dissociation. (Hartridge and Roughton.†)

In cases of CO poisoning there is no dyspnœa. This is because, as we have learned, the dyspnœic response to oxygen lack is a reflex from the carotid body when the oxygen *tension* of the arterial blood is lowered. But in CO poisoning the oxygen tension of the blood is normal, though its oxygen content is reduced. The other symptoms of oxygen lack are, however, more serious than the reduction of the oxygen saturation would seem to warrant. This is because CO alters the shape of the oxygen dissociation curve, removing the S-shape, and moving it to the left, so that the blood less readily gives off oxygen to the tissues.‡

Many other gases have specially poisonous properties.

Irrespirable gases are those which are so irritating that they produce spasm of the glottis. Such are ammonia, chlorine, sulphur dioxide, a high percentage (*e.g.* 20 per cent.) of CO_2, and many others.

PROTECTION OF THE LUNGS FROM DUST. Although the inhaled air contains dust particles, very few of these reach the deeper structures of the lung, and those that do so are quickly removed. Coarser dusts are normally removed from the air as this passes over the extensive nasal mucosa, which is always kept moist with sticky mucus. In mouth-breathing this line of protection is greatly weakened, but the trachea and bronchial tubes are similarly moistened and so most of the dust particles reaching them are trapped. These are then carried upwards by ciliary action in the coat of mucus, and, reaching the mouth, are swallowed. This mechanism is remarkably efficient and very few dry dust particles ever reach the alveoli. Dust particles trapped in the larger tubes are carried upwards at a speed of 1–3 c.m./min.§ It has been claimed that the removal of dust-laden mucus is expedited by peristaltic contractions of the bronchial tubes synchronising with respiratory movements.

Extremely fine and scanty particles do, in spite of these defence measures, find their way into the alveoli. These are removed by wandering cells and deposited in the lymph glands at the root of the lungs. In town-dwellers and in coal workers these glands are always found to be black and to contain many cells that have carried ingested dust particles.

For Reference

CATCHPOLE and GERSH (1947). Pathogenetic Factors and Pathological Consequences of Decompression Sickness. *Physiol. Rev.*, **27**, 360.

ROBERTSON (1941). Phagocytosis of Foreign Material in the Lung. *Physiol. Rev.*, **21**, 112.

VENTILATION. A point of practical importance is the securing to each individual of sufficient fresh air. It is found that a dwelling-room becomes unpleasant and stuffy when the amount of CO_2 present has reached 0·1 per cent. This stuffiness is due to exhalations from the skin, lungs and alimentary canal. As these cannot be measured, it is taken as a cardinal rule in ventilation that the amount of CO_2 should not rise above 0·1 per cent.

An adult man gives off about 0·6 c. ft. of CO_2 every hour. Hence, in that time he raises the amount of CO_2 in 1,000 c. ft. of air from 0·04 per cent. (the normal amount in the atmosphere) to 0·1 per cent. He must, therefore, be supplied with 2,000 c. ft. of air per hour, in order to keep the amount of CO_2 down to 0·07 per cent. (Ordinary air contains 0·04 per cent. CO_2, therefore 2,000 c. ft. would contain 0·8 c. ft. CO_2, which, with the 0·6 c. ft. given off by the man, would be 1·4 c. ft., which is 0·07 per cent.). In crowded rooms the CO_2 may rise to 2 per cent. or more. The maximum rise in per cent. CO_2 after a considerable period is given approximately by the formula $\frac{70}{R}$, where $R =$ ventilation rate in c. ft./hr. per person.

---

* DOUGLAS, HALDANE and HALDANE. *J. Physiol.*, 1912, **44**, 275.

† HARTRIDGE and ROUGHTON. *Proc. Roy. Soc.*, 1923, **102A**, 595; **104A**, 376, 395; 1925, **107A**, 654; 1923, **94B**, 336.

‡ J. B. S. HALDANE. *J. Physiol.*, 1912, **45**, 22 P.

§ FLOREY, CARLETON and WELLS. *B. J. Exp. Pathol.*, 1932, **13**, 269; BARCLAY, FRANKLIN and MACBETH. *Amer. J. Rœnt. and Radium Therap.*, 1938, **39**, 673.

In order that the air may be easily renewed without giving rise to excessive draughts, a certain amount of cubic space must be allotted per person. Each adult in a room should have 1,000 c. ft. of space, and be supplied every hour with 2,000 to 3,000 c. ft. of air.

For Reference

Bainbridge (1931) (Edited by Bock and Dill). The Physiology of Muscular Exercise. (London : Longmans Green.)

Barcroft (1925 and 1928). The Respiratory Function of the Blood, I. and II. (Cambridge University Press.)

Cordier and Heymans (1935). Le centre respiratoire. (Hermann, Paris.)

Gesell (1925). The Chemical Regulation of Respiration. *Physiol. Rev.*, **5**, 551.

Haldane, J. S., and Priestley (1935). Respiration. (Oxford : Clarendon Press.)

Haldane (1927). Acclimatisation to High Altitudes. *Physiol. Rev.*, **7**, 363.

Henderson (1925). Physiological Regulation of the Acid-base Balance of the Blood. *Physiol. Rev.*, **5**, 131.

Killick (1940). Carbon Monoxide Anoxæmia. *Physiol. Rev.*, **20**, 313.

Schneider (1921). Physiological Effects of Altitude. *Physiol. Rev.*, **1**, 631.

Schmidt and Comroe (1940). Functions of the Carotid and Aortic Bodies. *Physiol. Rev.*, **20**, 115.

BOOK VI

THE INTAKE OF MATERIALS—NUTRITION, METABOLISM

HISTORICAL NOTE

ANCIENT and mediæval physicians, confounded by the rapidity with which food entering the alimentary canal became altered in appearance and reduced in amount, and unrestrained by any notions regarding the indestructibility of matter, did as would be expected, and cloaked their ignorance in words and doctrines. One speculation, due initially to ERASISTRATUS (*circa* 300 B.C.), and adopted with amplifications by GALEN (*circa* A.D. 180), was that the food was altered by a change called "concoction" into a principle, which, being absorbed and proceeding to the liver, was there converted into blood and endowed with "natural spirits": subsequently, as we have seen in earlier chapters, the vital spirits and animal spirits arose from it. This doctrine, being based on authority rather than on fact, died hard, or at least late.

The first approach to a realistic view of the nature of digestion was made by VAN HELMONT (1577–1644), who no doubt fancying a resemblance between the familiar phenomena of alcoholic fermentation and the vulgar facts of flatulence, saw the course of digestion as a series of fermentations, the ferment of the stomach being prepared in the spleen, and mixed with acid. Since the bile was the only then known juice poured into the alimentary canal, his guess was more astute than he knew: but we should not read too much prophecy into it, seeing that the biological, and still more the enzymic, view of fermentation lay centuries ahead. He did, however, believe that the tissues took up substances from the blood.

The notion of glands, secretions, and ducts having been put forward by STENSEN, who discovered the parotid duct in 1662, it was natural for FRANCISCUS SYLVIUS (1663) to follow up the views of Van Helmont, and to regard the changes of digestion as being what would now be termed chemical in nature, with the digestive juices, especially saliva, as active agents.

At about that time (1664) DE GRAAF performed the surprising operation of cannulating the pancreatic duct of a dog, and so collected pancreatic juice: its functions were not known, however, till long afterwards.

The science of physics was then beginning to take shape, and the new knowledge was regarded as the key to all natural phenomena. So, due to the influence of the iatro-physical school, led by BORELLI (1608–79), digestion, to be in the fashion, came to be considered as a progressively finer and finer trituration, in which the act of mastication was the first and crudest step. But the balance swung back again when the wide genius of BOERHAAVE, in 1708, found in compromise the answer which is still held as correct, viz. that both mechanical and chemical processes play their parts in reducing the food to such a state as enables it to be absorbed into the blood. In the absence of both chemical knowledge and directed experiment, however, speculation regarding the nature of the chemical changes still continued, and even the immortal HALLER (*circa* 1760) got no further than regarding them as akin to putrefaction.

Instructive experiments began with the observations of RÉAUMER (1752). He induced a tame kite to swallow perforated metal tubes containing various food materials, which were regurgitated to order after given intervals of time. Meat and bone had undergone no putrefaction, but, although protected from mechanical assault, were soaked with liquid and partly dissolved, while starchy foods were unaffected. His conclusion that the liquid had a solvent action was confirmed by SPALLANZANI (1777), who placed bits of sponge in the tubes, and showed that the liquid squeezed from them would digest meat *in vitro* if kept warm. Spallanzani, and, quite independently, STEVENS of Edinburgh (1777) applied these methods to man, and confirmed the results, but failed to note the acidity of the gastric juice, the first inkling of this being given by CARMINATI in 1785, the final identification of HCl by PROUT (1834). For a time, all the actions were attributed to the acid, although EBERLE had, in 1834, made an active artificial juice by digesting gastric mucosa with dilute HCl. A classical investigation, the first on a gastric fistula, was made by WILLIAM BEAUMONT, a United States army surgeon, between 1825 and 1833. The subject, Alexis St. Martin, a Canadian voyageur, received a gunshot wound which, after healing, left a fistulous opening into the stomach. Through this orifice Beaumont observed the secretion of gastric juice, the state of the empty stomach, its movements, etc., and noted the times taken to digest foods of various types. Distillation of the gastric juice which he collected confirmed the presence of considerable amounts of HCl. Experimental gastric fistulæ in animals were first made in 1842 by Bassow, and have since been much improved, notably by Pavlov.

Belief in the existence of specific "ferments" still existed, but the first to be named was pepsin, by Schwann in 1835. New enzymes were soon added, *e.g.* ptyalin by Miahle (1845); two enzymes of the pancreatic juice were described by Claude Bernard in 1856 and trypsin by Kühne in 1876; enterokinase was discovered by Pavlov in 1899, etc.: the generic term "enzyme" was introduced by Kühne in 1878, the catalytic nature of their action having been previously pointed out by Berzelius in 1838.

Physiologists have always tended to evolve general theories of the secretory process, and these theories have closely reflected the state of contemporary knowledge. In the 18th century, there being little knowledge of the minute structure of glands, the opinion held (Haller, 1760) was of a direct communication between the gland substance and the finest extremities of the blood-vessels, through which a kind of filtration of plasma took place. By the mid-19th century, a sound morphological knowledge existed: Johannes Müller (1830), held that the secretion was formed by the walls of the acini, the constituents passing from the blood of the encompassing capillary network, in amount and kind depending on the nature of the living acinar cells. After the enunciation of the cell theory (1838), views not unlike those now held, such as the elaboration of special substances (*e.g.* Kolliker (1850)) became current. The influence of physics also made itself felt: out of the membrane diffusion studies of Dutrochet (1827), and Graham (1850), arose the mechanical theory of Ludwig (1856), which was countered by Heidenhain (1883), who showed that physical causes alone provided an inadequate explanation.

The first indication of the means by which the secreting glands were brought into action was provided when Ludwig, in 1851, showed that the submaxillary gland secreted if the peripheral end of the chorda tympani was stimulated: secretory nerves to other glands were expected, and were, in fact, discovered later, *e.g.* the vagus to the pancreas and stomach. Bayliss and Starling, in 1902, however, showed that introduction of dilute acid into the duodenum provoked secretion of pancreatic juice by even the denervated pancreas: this led them at once to the discovery of secretin, and to the conception of humoral stimuli, to which they gave the general name of hormones.

Movements of the stomach were observed by Beaumont, and those of the intestines, familiar from ancient times, were systematically described by Ludwig in 1861. By the use of graphic methods Bayliss and Starling (1899) added considerably to our knowledge of the intestinal movements; the first use of X-rays, with opaque meals, now so extensively used in the clinic, was due to Cannon, who studied the movements of the stomach in 1898, and of the intestine in 1902.

During the latter part of the 19th century, there was considerable expansion in detail of the knowledge of the general course of digestion, of the secretion of the digestive juices, of absorption, and of the subject of physiological chemistry. For example, Kühne's work between 1876 and 1883 on the course of pancreatic digestion, of which leucine and tyrosine were recognised as products, Heidenhain's studies of the structure of the gastric mucosa, and introduction of the gastric pouch in 1879 (subsequently improved by Pavlov in 1898), and a series of important investigations by Pavlov from 1888–1900 on the work of the digestive glands.

Until 1828 the opinion was held that organic substances, which were by then being isolated in increasing numbers, could only be manufactured by living organisms, but in that year Wöhler showed that urea could be made synthetically. This paved the way for investigations on the constituents of the body, and of foods, which was crowned by the classical work of Fischer on the carbohydrates, purine bodies, and proteins. The idea of chemical material alterations in the body, now called metabolism, was introduced about 1836, and an attempt to make up a material balance sheet for the body was first made by Boussingault in 1836. The development of thermodynamics in the mid-19th century opened up the possibility of studying an energy balance sheet; the introduction of a respiration apparatus by Regnault and Reiset in 1849, its improvement by Voit in 1861, and the classical studies on heat values and the calorific balance sheet by Rubner about 1891, and by Atwater and Benedict from 1897 onwards, placed that subject in the catalogue of accepted scientific facts. The indispensability of protein in the diet was shown by Majendie (1819), and the conception of nitrogenous equilibrium was due to Voit (1857).

Lavoisier (1789) believed that in the lungs, oxygen directly oxidised carbon and hydrogen: Liebig (1842) taught that it was protein, carbohydrate and fat that were oxidised, the protein being only oxidisable in muscle, and during its contraction. Out of these various researches there grew the opinion that appropriate amounts of protein, carbohydrate, fat, salts and water were all that were needed to constitute a complete diet. The accessory food factors came to be recognised as a result of the work of Lunin (1880), Eijkman (1897), Holst and Fröhloch (1907), Hopkins (1912), and Funk (1911), the name "vitamin" being proposed by the last-named. The constitution of most of the known vitamins has been worked out since 1930, and several of them have been made synthetically.

The unfolding of the problems of intermediary metabolism has been largely linked up with the recent rapid development of biochemistry. One of the most important studies was that of Knoop (1910) who traced the oxidation products of amino-acids by linking them with a benzene ring. Recent developments which have proved fruitful are the use of the tissue slice method, conjoined with respirometric and microchemical technique, and the employment of compounds containing isotopic elements as a means of tracing intermediary products in metabolism. The development of rapid and accurate methods of microchemical analysis has also provided biochemists with most valuable tools, and to these has been added the employment of chemical substances having specific effects on particular tissues, or on particular enzymes : drugs, *e.g.* like atropine, acetyl choline, iodoacetate, etc. Probably no drug has played a larger part in the history of the metabolism of carbohydrate, and, incidentally, that of other materials, than has phlorrhizin (v. MERING, 1885). The relation of the pancreas to diabetes mellitus was discovered by v. MERING and MINKOWSKI (1889), the presence of sugar in diabetic urine and blood having been long known. The then hypothetical anti-diabetic hormone of the pancreas was named insulin by SHARPEY-SCHAFER in 1916, and, after being sought by many, was finally extracted from pancreas, and isolated, by BANTING and BEST in 1922.

The nature of the chemical changes by which the ultimate oxidation of the food materials was effected has been a fruitful source of experiment and of speculation. Before about 1920, oxidation was believed to be due to a direct combination between oxygen and the compound to be oxidised, though it was doubtful as to how this might be brought about in animal tissues. About that time however, WIELAND and THUNBERG showed that certain substances, such as succinic acid, could be oxidised in complete absence of free oxygen, by transfer of two hydrogen atoms to a reducible substance under the influence of tissue enzymes called dehydrogenases. Almost simultaneously Warburg put forward the view that oxidation can be brought about by iron-containing catalysts that activate free oxygen. By about 1935 it was realised from KEILIN'S work that both mechanisms were employed by the tissues, dehydrogenation reactions, often in many stages, leading finally to oxygen activation by cytochrome oxidase at the end of the chain.

The means by which carbon dioxide is finally produced as an end-result of oxidative metabolism, and in which the tricarboxylic acid cycle plays an important role, began to be realised about 1940 from the work of KREBS, SZENT-GYORGYI and others.

CHAPTER XLI

THE EXCHANGES OF MATTER AND ENERGY IN THE BODY

All the energy which leaves the body as heat or work is derived from processes of oxidation, the carbon, hydrogen, nitrogen and sulphur of the foodstuffs being oxidised and eliminated in the form of carbon dioxide, water, urea, sulphates, etc. In a starving animal this is associated with a loss of body substance. The necessity for taking food is determined by the need of replacing this loss. Most food-stuffs cannot be utilised directly as a source of energy, but must be altered and built up into the living protoplasm. The total amount of living material in the body, though fairly constant in the adult, may yet undergo alterations under some conditions, and especially in the growing animal. We have in this chapter to inquire into :

(1) The relation between the total amount of material taken up by the body and the total amount given out—the *material balance sheet ;*

(2) The relation between the potential energy available to the body and the amount of energy set free as heat or work—the *energy balance sheet ;*

(3) The nature and amounts of the foodstuffs necessary for maintaining the weight of the body constant or providing for its growth and health ; and

(4) The significance of the various classes of foodstuffs as sources of energy and in the replacing of tissue waste.

1. THE MATERIAL BALANCE SHEET

METHODS EMPLOYED

Income. The determination of the material exchanges of the body involves comparison of its income and output. The income consists of the foodstuffs and oxygen. The foodstuffs may be divided into two classes, namely (1) the organic foodstuffs, which on oxidation may serve as sources of energy, and (2) the inorganic foodstuffs, such as salts and water.

The latter class neither add to nor subtract from the total energy of the organism, but their presence is essential ; and as they are contained in the various excreta, a corresponding amount must be present in the food to make good this loss.

In spite of the variety of foods, their essential organic constituents can be assigned to the three classes, proteins, fats and carbohydrates ; and analyses of the food must give the relative amounts present in digestible form of these three proximate principles.

In the proximate analysis of the foodstuffs the water is determined by drying at 105° C., the inorganic constituents by ashing the dried remainder. The nitrogen (Kjeldahl's method), multiplied by 6·25, gives approximately the total protein. The fat is determined by extracting the dried food in a Soxhlet apparatus with ether. Carbohydrates may be determined by boiling the food with dilute acids in order to convert all carbohydrate into hexoses, which are then estimated as glucose. The carbon and hydrogen content of the food can easily be calculated.

In order to attain greater accuracy, some observers have made a complete food, which is prepared in large quantities at the beginning of the experiment and used as the sole diet throughout. The monotony of such a diet tends to interfere with the success of the experiment, since the subject loses his appetite.

For approximate calculations, especially of dietaries of whole communities, the labour of analyses is lightened by the fact that nearly all the ordinary foodstuffs have been analysed and their average composition published.*

---

* R. H. A. Plimmer, "Analysis and Energy Values of Foods," H.M. Stationery Office, 1921. Atwater and Bryant, *Bull.* 28, U.S. Dept. Agricult., 1906. McCance and Widdowson, "Chemical Composition of Foods" (M.R.C., Rep. No. 235, 1940).

Output. The material output of the body is turned out by the kidneys, the alimentary canal, the lungs and the skin. These excreta must therefore be collected and analysed. In most cases the losses by the shedding of the cuticle, hairs and nails, and so on, are so small that they may be disregarded.

The nitrogen of the food, and that derived from the disintegration of the tissues of the body, is excreted almost exclusively in the urine, but a small amount is thrown out by the alimentary canal. The total nitrogen must therefore be determined both in the fæces and in the urine. The nitrogen in the fæces is derived from two sources. Part represents those nitrogenous constituents of the diet which have resisted digestion, and a certain amount is derived from the intestine itself. Even during complete starvation fæcal matter is formed in the intestine, and it has been calculated that normally about 1 g. of nitrogen a day is excreted in the secretions entering the gut and contributes to the formation of the fæces. It is usual, therefore, to regard 1 g. of the nitrogen of the fæces as belonging to the *output* of the body, while the balance belonging to undigested foodstuffs is subtracted from the total nitrogen of the latter in reckoning the real *income* of the body. A small amount of nitrogen is also lost by sweat, but this can usually be disregarded.

If an animal is kept on a sufficient diet which contains a regular amount of nitrogen, after a few days a condition known as *nitrogenous equilibrium* is set up, *i.e.* the total nitrogen of fæces and urine is exactly equal to the total nitrogen of the food. The same applies to the sulphur, as shown in the following Table (quoted by Tigerstedt):

Days of experiment.	Nitrogen of food.	Nitrogen excreted.	Per cent. difference.	Sulphur ingested.	Sulphur excreted.
1–7	154·81	153·02	— 0·51	—	—
8–17	213·72	213·26	— 0·21	12·77	12·79

In order to express the nitrogenous metabolism in terms of protein, we multiply the total nitrogen of the excreta by 6·25. This will give the total protein which has been broken down during the period of the experiment.

Respiratory Exchange. This gives us the oxygen intake and the carbon dioxide output. The estimation of these involves the use of some form of respiration apparatus. Water output in the breath can be calculated, or determined directly.

The methods which have been used for this purpose are very numerous, but are of two main types, the closed or re-breathing type and the open or expired-air-analysis type. In the closed type, the subject breathes an atmosphere which circulates in a closed apparatus; the carbon dioxide produced is absorbed, and can be weighed, and oxygen is added in measured amount to keep the volume of the enclosed atmosphere constant: the nitrogen is breathed over and over again. In the open type fresh air is inspired, and the expired air is subjected to measurement and analysis in one of several ways. The Benedict apparatus is on the closed principle; Haldane's and Douglas' methods are on the open principle.

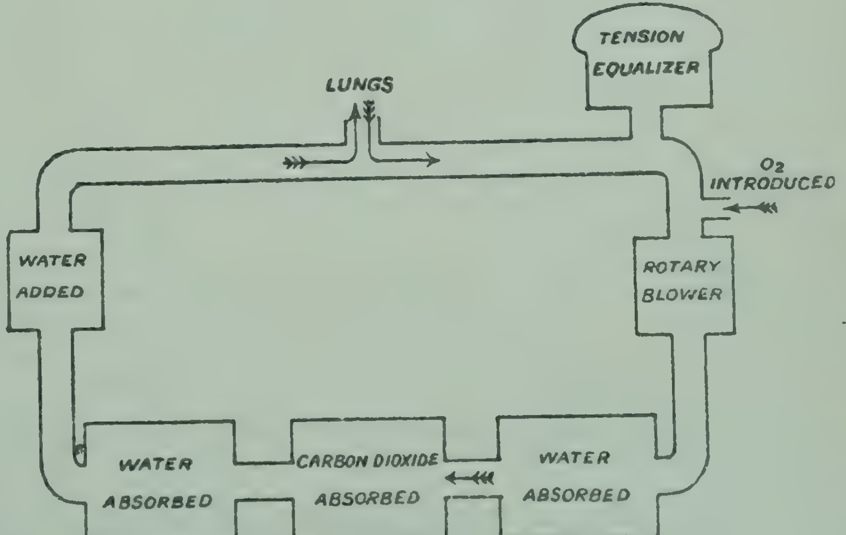


Fig. 522. Air Circuit in Benedict's Respiration Apparatus.

I. The Method of Benedict. The subject of investigation is placed in a closed chamber containing a given volume of air. The carbon dioxide produced is absorbed in caustic alkali, and the oxygen consumed is made good by allowing oxygen in measured amount to flow into the chamber. The inflow of oxygen is regulated so as to keep the pressure of air in the chamber constant. At the end of the experiment, the alkali is weighed to find the amount of carbon dioxide absorbed ; or the CO_2 can be expelled by acid and its volume measured direct. The air in the chamber is also analysed at the end, to correct for any change.

The Atwater-Benedict Respiration Calorimeter (Fig. 526, p. 792) is on a similar principle.

In a modification of this method * applicable to clinical purposes, the subject breathes through a nose-piece into a wide metal tube, the mouth being kept closed. The metal tube forms part of a closed system through which a current of air is maintained by means of a blower. In the course of the current of air are interposed vessels for the absorption of carbon dioxide and of water, and the volume of gas in the system is maintained constant by admitting oxygen to it as the oxygen is used up in respiration. The tension equaliser (Fig. 522) consists of a pan closed by a loose rubber bag. The whole circuit to commence with

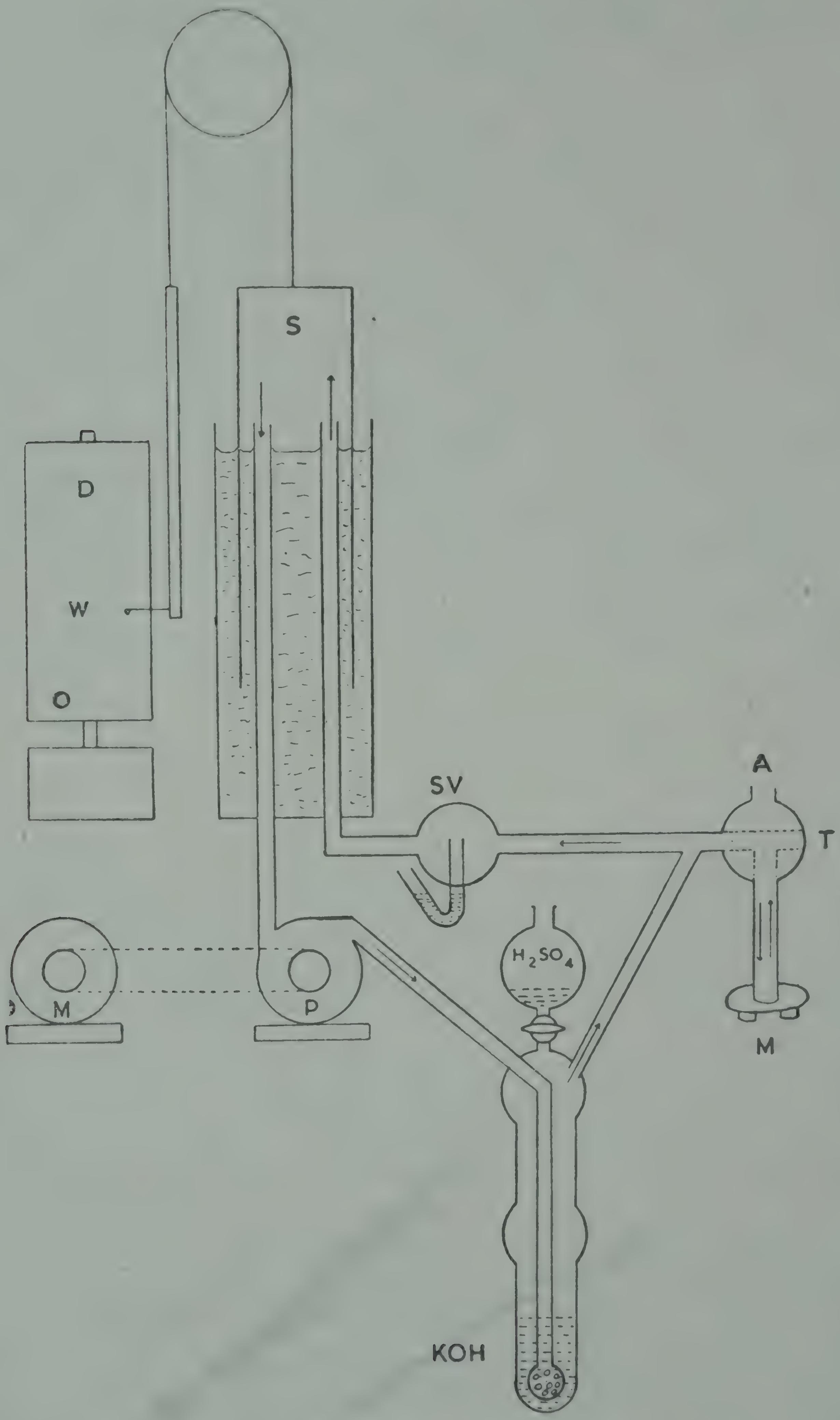


FIG. 523. Knipping Recording Spirometer. (C. F. Palmer & Co.)

---

* BENEDICT. *Boston Med. & Surg. J.*, 1918, **178**, 667 ; 1920, **183**, 449.

is filled with oxygen-enriched air, slightly filling out the rubber bag. As air is drawn into the lungs the rubber diaphragm sinks, to rise again with expiration. The respiratory movements can thus proceed without appreciably altering the pressure within the closed system of tubes. By the admission of oxygen (measured through a meter) the volume is adjusted so as to keep the bag from becoming either too distended or too flattened. As air leaves the lungs and is swept along by the blower, it flows through bottles containing sulphuric acid for the removal of water vapour and then through a cylinder, filled with soda lime for the absorption of carbon dioxide. From here it passes again through sulphuric acid for the absorption of water given off by the soda lime. Since the dry air would be uncomfortable to breathe, it is then carried through another bottle containing water with a trace of sodium carbonate for the neutralisation of any acid fumes given off by the sulphuric acid. It then passes back to the tube from which the subject is breathing. In this way it is possible to determine very accurately the amount of oxygen used up, and the amount of carbon dioxide given off, in the course of an experiment lasting one to three hours, or longer.

II. The Recording Spirometer. The Knipping apparatus (Fig. 523) is a form which enables CO_2 production also to be measured. The air is circulated by a blower P; the subject breathes through the mouthpiece M and the CO_2 is absorbed in KOH solution. The descent of the spirometer bell S gives a measure of the oxygen usage. As oxygen is used, the respiratory record is written on a sloping line; more oxygen is added through a gas meter from time to time to restore the respiratory base line. The rate of oxygen usage at any time is given by the slope of the respiratory tracing. After the oxygen usage has been determined the tap T is closed and the CO_2 liberated from the KOH by adding acid from the tap-funnel. After cooling, the volume increase represents the CO_2 liberated.

III. The Method of Haldane.* This method is convenient when dealing with the gaseous exchanges of small animals, such as mice, rats, or frogs. The animal is placed in the chamber C (Fig. 524). On the inlet side of the bottle is a series of bottles of sulphuric acid and soda lime. On the outlet side is a corresponding series of vessels for the absorption of water and of carbon dioxide. Air is sucked through the whole apparatus by means of a water pump. The animal is thus supplied with dry air freed from carbon dioxide. All water or carbon dioxide produced by the animal is absorbed by the vessels interposed on the course of the outgoing air. These vessels are weighed at the beginning and at the end of the experiment, and the difference in weights will therefore give the amounts of carbon dioxide and water which have been discharged by the animal.

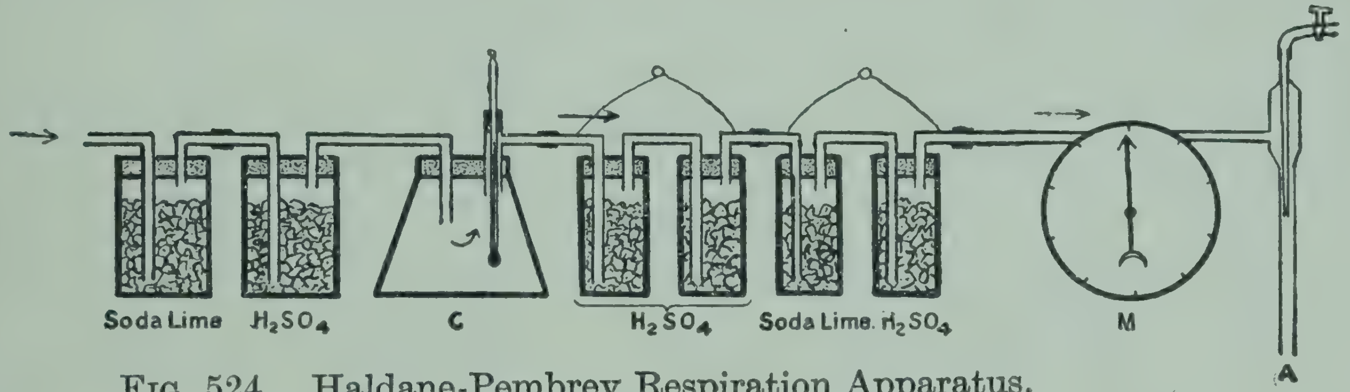


FIG. 524. Haldane-Pembrey Respiration Apparatus.
C. Chamber for animal. M. Gas meter.

The intake of oxygen is determined indirectly. Since it gives off only carbon dioxide and water, and absorbs only oxygen during its stay in the chamber, the loss of weight of the animal during its stay in the chamber, subtracted from the total weight of carbon dioxide plus water it gives off, will represent the weight of oxygen absorbed.

IV. The Douglas Bag. A convenient method for estimating the respiratory exchanges of man is by the use of the Douglas Bag, by which the expired air is collected (*v.* p. 723). When a sufficient volume of air has been collected, the bag is removed, the time taken in collecting being accurately noted. The amount of air expired during this time is measured by emptying the bag through a gas meter. A sample of its contents is analysed and the oxygen and CO_2 in it determined. From the analysis and measurement of the expired air we can calculate the respiratory metabolism during the time of the observation. This method has been used for determining the metabolism of soldiers in training, of industrial workers, athletes, etc.

Calculation of Results. In the closed circuit methods, the weight or volume of oxygen used and carbon dioxide produced are given directly, and any necessary corrections, *e.g.* of volumes to dry volumes at N.T.P., are obvious. With the open methods, some explanation is necessary. The subject inspires ordinary air which contains 20·96 volumes per cent. of oxygen, 79 per cent. N_2, and 0·04 per cent. CO_2, and it is the volume and composition of his expired air during a given interval which is actually determined. As regards CO_2, let us suppose that the expired air contains 4·10 per cent., *i.e.* an increase of 4·06 per cent.; then, if his volume of expired air is 8·5 l. per min., the volume of CO_2 expired is $85 \times 4{\cdot}06 = 345$ c.c. per minute. The oxygen is less direct because, as explained in Chapter XL., the expired air has generally a different, mostly a smaller, volume than the inspired, the *percentage* of nitrogen being raised owing to utilisation of oxygen which is not replaced by CO_2. The *total amount* of nitrogen breathed out is, however, equal to that inspired. Hence, if the expired air contains, say, 79·5 per cent. N_2 and 16·4 per cent. oxygen, then each 100 c.c. expired air corresponds to $\frac{100 \times 79\cdot5}{79} = 100\cdot6$ c.c. inspired air and this contained $\frac{20\cdot96 \times 100\cdot6}{100} = 21\cdot09$ c.c. oxygen, so that the oxygen utilised per minute was (21·09 − 16·4) × 85 = 398 c.c.

---

* HALDANE. *J. Physiol.*, 1892, **13**, 419.

The Respiratory Quotient is obtained, when required, from the respiratory exchange. It is the ratio of the volume of carbon dioxide produced during a given time to the volume of oxygen utilised during the same time. (Both gases should, of course, be measured at the same temperature and pressure.) In the above example it is $\frac{345}{398} = 0\cdot87$.

If the R.Q. is exactly 1·0, the percentage of nitrogen is unaltered. When the R.Q. is greater than unity, the percentage of nitrogen in the expired air is less than that in the inspired air, but the same type of calculation as that given above can be applied.

By these methods we may arrive at a correct idea of the total income and output of an individual for periods up to many days. The following details by Atwater and Benedict may serve as an example of the results obtained in such experiments :

SUMMARY OF INCOME AND OUTPUT OF NITROGEN, CARBON AND HYDROGEN IN MAN (ATWATER AND BENEDICT *)

State.	Total No. of Days.	Average Daily Income. (Grams.)	Average Daily Output. (Grams.)					Gain or Loss. (Grams).
		Food.	Fæces.	Urine.	Water.	Respn.	Total.	
Rest	42	NITROGEN. 18·5	1·1	17·8	—	—	18.9	−0·4
		CARBON. 253·3	9·9	12·4	—	217·9	240·2	+13·1
		HYDROGEN. 39·2	1·4	3·5	41·8	—	46·7	−7·5
		(Protein lost by body = 2·4 grams) (Calc. from N. loss)		(Fat gained by body = 19·2 grams) (Calc. from C. gain)				
Fast of 1 or 2 days	5	NITROGEN. 0·0 .	0·0	13·1	—	—	13·1	−13·1
		CARBON. 0·0 .	0·0	9·9	—	184·4	194·3	−194·3
		HYDROGEN. 0·0 .	0·0	2·7	98·3	—	101·0	−101·0
		(Protein lost = 82·0 grams) (Fat lost = 198·3 grams)						

In the resting experiments, in which the subject was on an ordinary diet, it is seen that the intake and output almost balance ; there was, however, a slight loss of 0·4 gramme of nitrogen, which means that about 2·4 grammes of protein from the body had been lost. Some fat had, however, been laid on, as shown by the gain of 13·1 grammes of carbon. In the fasting experiment the body, of course, lost considerably.

For Reference

RICHARDSON (1929). The Respiratory Quotient. *Physiol. Rev.*, **9**, 61.

Growth. In growing individuals, materials from the food are converted into body substances. But the proximate constituents of the diet are not retained in the same proportions as they occur in the diet. The following table shows that carbohydrate, though the most abundant constituent of the diet, hardly accumulates at all.

---

* ATWATER and BENEDICT, 1903. *U.S. Dept. Agric. Bull.*, 1903, *No.* **136**, p. 120.

AVERAGE COMPOSITION OF HUMAN BEINGS (YOUNG *)

	Weight of Material (kg.).			Percentage of Total.		
Material.	New Born.	Adult.	Added during Growth.	New Born.	Adult.	Added during Growth.
Total mass	2·9	70·0	67·1	—	—	—
Water	2·1	41·0	38·9	72	59	58
Dry matter	0·8	29·0	28·2	28	41	42
Carbohydrate	0·01	0·40	0·39	0·35	0·57	0·58
Fat	0·35	12·60	12·25	12·1	18·0	18·2
Protein	0·34	13·10	12·76	11·7	18·7	19·0
Salts	0·10	3·00	2·90	3·5	4·3	4·3

This must mean that carbohydrate is either burnt as fuel, or is converted into some other constituent, such as fat.

Of the total food taken during the years of growth, it will be seen from the following table, that only 0·02 per cent. of carbohydrate can be regarded as being retained, whereas the amounts of protein and fat respectively retained amount to over a hundred times that amount.

RATE OF ADDITION OF MATERIAL ADDED TO HUMAN BODY DURING 7,000 DAYS (APPROX. 19 YEARS) OF GROWTH (YOUNG *)

Material.	Weight of Material in Adult Body (kg.).	Weight of Material added during Growth (kg.).	Rate of Addition of Tissue (g./day).	Approx. Percentage of Ingested Food retained in Body during Growth.†
Total mass	70·0	67·1	9·6	0·32
Water	41·0	38·9	5·6	0·22
Dry matter	29·0	28·2	4·0	0·85
Carbohydrate	0·40	0·39	0·06	0·02
Fat	12·60	12·25	1·7	2·27
Protein	13·10	12·76	1·8	2·40
Salts	3·00	2·90	0·4	2·00

These figures indicate that growing children require, for their size, more protein and fat than adults, but that increased amounts of carbohydrate, except in so far as it can form fat, are less urgently needed.

2. THE ENERGY BALANCE SHEET OF THE BODY

Income. The energy income of the body is measured by the potential energy of the foodstuffs, *i.e.* the amount of energy which can be evolved, either as heat, work, or in any other form, by the oxidation of the foodstuffs to the end products which occur in the body. We express the total potential energy of a foodstuff in *Calories*. The heat value of any given food is the number of Calories ‡ which it evolves on complete combustion with oxygen, and is determined by burning a weighed quantity of the dried foodstuff in oxygen in a bomb calorimeter.

Instead of using the bomb calorimeter, the heat of combustion can be determined *indirectly* by finding how much oxygen is needed to burn the foodstuff. As seen from the table on p. 791, for common foodstuffs the Calories

* YOUNG. *Sci. Prog.* 1948,, **141,** 13.

† Based on mean daily food intake for 0–19 years of 300 g. carbohydrate, 75 g. fat, 75 g. protein, 20 g. salts, 2600 ml. water (including water formed by oxidation).

‡ A calorie is the amount of heat necessary to raise a gramme of water from 15° C. to 16° C. A kilocalorie, or large Calorie (printed with a capital C) is equal to 1000 calories and is the unit employed in most physiological work.

yielded per litre of oxygen used is roughly 4·8. The determination is carried out in a modified closed-circuit apparatus, called an oxy-calorimeter, in which the substance is burned in a circulating current of oxygen. A suitable apparatus is shown in Fig. 525.

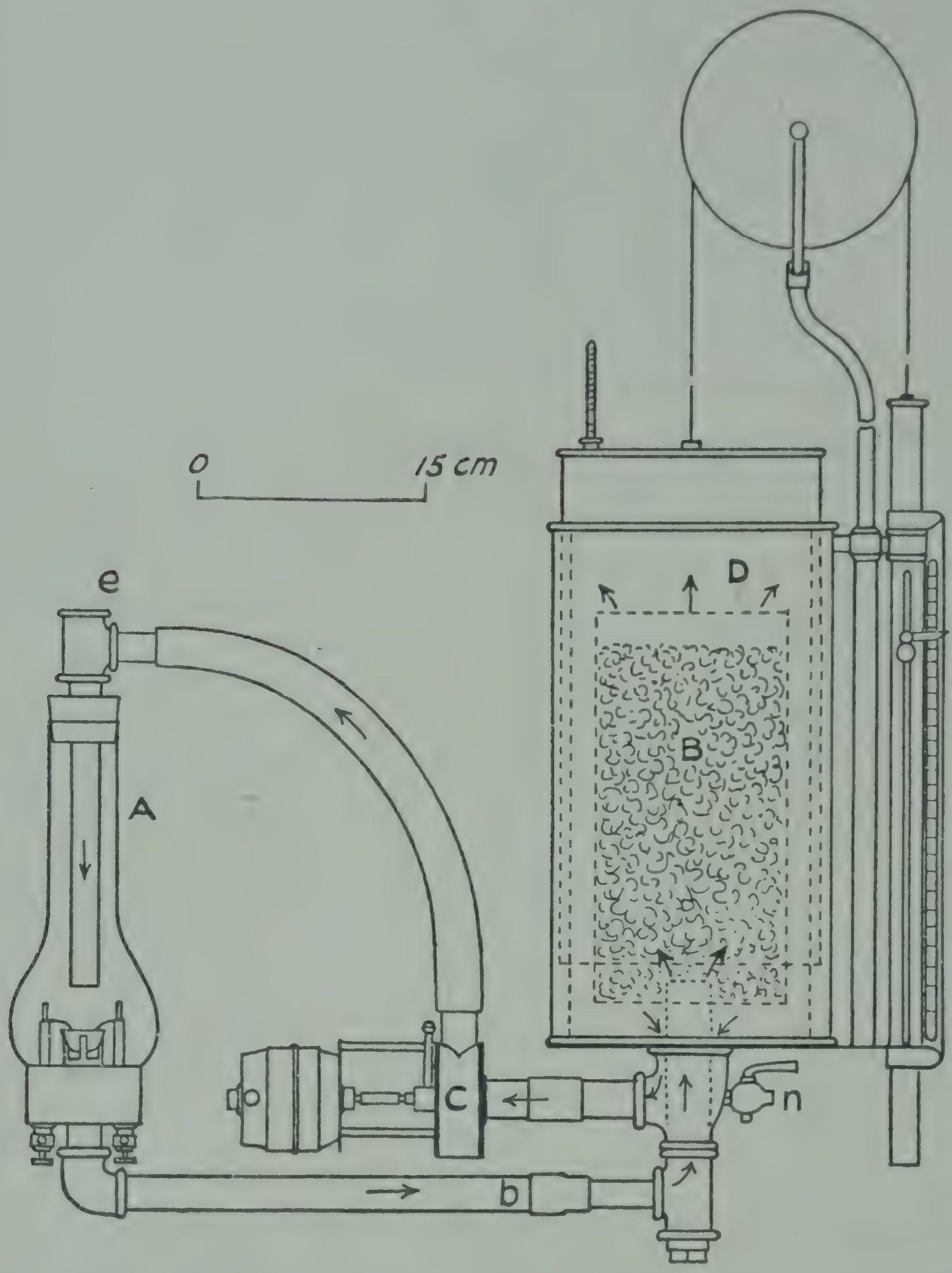


FIG. 525. Simplified oxy-calorimeter for determining energy values of foods, foodstuffs and excreta. (BENEDICT and FOX, *Journ. Biol. Chem.*, 1925, **66,** 783.)

The substance to be tested is dried and weighed into a small nickel crucible, inside the lamp-glass A. The apparatus is filled with oxygen, circulated by the rotatory blower C. The spirometer bell D gives the variation in volume, and is filled with soda lime to absorb CO_2. Combustion is started by electrical heating of the substance, and, after completion and cooling, the oxygen used is read off. Except when very accurate results are needed, an exact analysis of the substance is not necessary; the average calorific values of oxygen may be taken to be: for nitrogen-rich substances, 4·68 Calories per litre of oxygen; for fats, 4·7 Calories; for carbohydrate-rich foods, 5·0 Calories; for a mixed diet of carbohydrates and fats, 4·825 Calories; and for fæces, 5·0 Calories.

In the case of proteins it is necessary to draw a distinction between the heat value on combustion, and the physiological heat value. Since carbohydrates and fats undergo complete oxidation in the body to carbon dioxide and water the values of these foodstuffs to the organism are identical with their absolute heat values. When proteins are burnt, however, the nitrogen is converted into various gases. In the animal body, no nitrogen is eliminated in the gaseous form, but it is excreted as urea and allied substances still endowed with potential energy, which can be set free when their oxidation is completed by burning.

The following heat values have been obtained for different foodstuffs :

GASEOUS AND ENERGY RELATIONSHIP IN THE COMBUSTION OF SOME COMMONLY METABOLISED COMPOUNDS. (BENEDICT and FOX, *Journ. Biol. Chem.*, 1925, **66**, 783.)

Substance.	Oxygen required to oxidise 1 g.	Produced in the oxidation of 1 g.		Calories per litre of oxygen.
		Carbon dioxide.	Heat.	
	c.c.	c.c.	cals.	
Starch	829·3	829·3	4·20	5·06
Cane sugar	785·5	785·5	3·96	5·04
Glucose	746·2	746·2	3·74	5·01
Lactic acid	745·9	746·0	3·62	4·85
Animal fat	2013·2	1431·1	9·50	4·72
Human fat	1990·8	1420·4	9·54	4·79
Protein	956·9	773·8	4·40	4·60
Acetone	1542·9	1157·2	7·43	4·82
β-Oxybutyric acid	968·2	860·7	4·69	4·85
Ethyl alcohol	1459·5	972·9	7·08	4·85

In order to determine the physiological heat value of protein, we must subtract from its absolute heat value the heat value of the excretory products as they leave the body. The physiological heat value of proteins was determined by Rubner in this way : A dog was fed with a protein of absolute heat value, of 5·345 Cal./g., and its urine was collected, dried, and its heat value determined by combustion. It was found that, for each g. of protein which had undergone disintegration in the body, urine was passed of a heat value of 1·0945 Calories. The heat value of the fæces formed under the same diet was 0·1845 Calorie per g. protein. Rubner further reckoned the heat required for the solution of the proteins and of the urea, at 0·05 Calorie. The reduced or physiological heat value of protein is therefore equal to $5{\cdot}345 - (1{\cdot}0945 + 0{\cdot}1854 + 0{\cdot}05) = 4{\cdot}015$ Calories.

It is usually unnecessary to reckon out accurately the relative values of the different kinds of protein, carbohydrate, etc., contained in each diet, and Rubner calculated the *average* physiological heat values of the three classes of foodstuffs, and found them as follows :

1 grm. protein = 4·1 Calories
1 grm. fat = 9·3 ,,
1 grm. carbohydrate = 4·1 ,,

These figures are accurate only for a diet containing the normal proportion of vegetable to animal foods—60 to 40. The heat value of vegetable protein is as a rule less than that of animal protein.

Output. The energy expenditure of the body is measured in Calories and can be determined either directly by placing the subject in a suitable calorimeter, or indirectly from the oxygen consumption.

THE RESPIRATION CALORIMETER. The Atwater-Benedict * respiration calorimeter consists of a room or chamber with double non-conducting walls. All round the inner wall of the room are fitted coils of pipes through which a stream of water flows. The pipes are fitted with discs in order rapidly to take up heat produced in the room. The current of water is accurately adjusted so as to maintain the temperature of the inner wall constant. As the inner wall and outer wall are kept at the same temperature, no heat is lost to the exterior, the whole of the heat produced by the animal or individual in the chamber being communicated to the water passing through the chamber. The temperatures of the entering and leaving water are taken by thermometers readin to a thousandth of a degree Centigrade. Knowing the amount of water tha has passed through in a given time and the difference in temperature durin the same time, it is easy to calculate the amount of heat given off by the anima under investigation. It is generally convenient to maintain a constant difference of temperature between the entering and leaving water, by appropriat adjustment of the amount of water passing through the apparatus. Th equality of temperature between the inner and outer casing is recorded b electric thermo-couples, any difference of temperature being at once compensated by electrically warming the cooler part. The chamber contains a

---

* ATWATER and BENEDICT. *A Respiration Calorimeter. Publ. No.* 42, *Carnegie Inst., Washington*, 1905.

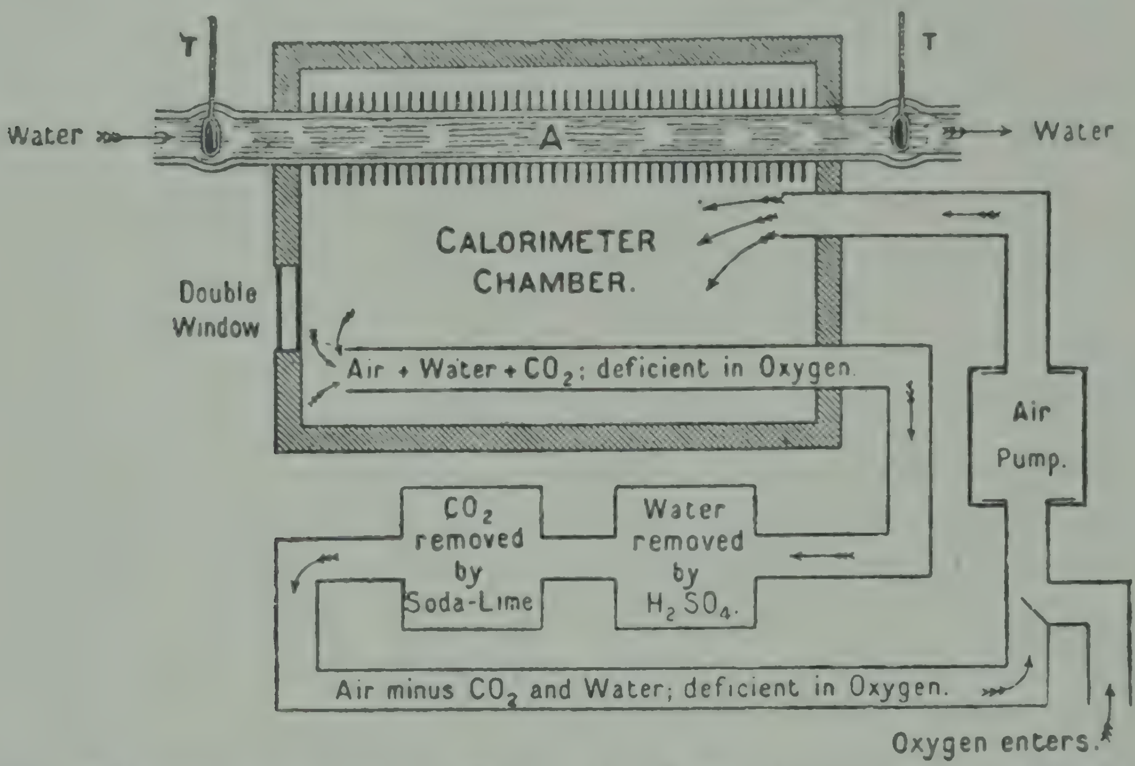

T
T
Water
A
Water
CALORIMETER CHAMBER.
Double Window
Air + Water + CO_2: deficient in Oxygen.
Air Pump.
CO_2 removed by Soda-Lime
Water removed by H_2SO_4.
Air minus CO_2 and Water; deficient in Oxygen.
Oxygen enters.

FIG. 526. Diagram to show the Principle of the Atwater-Benedict Calorimeter. (After HALLIBURTON.)

bicycle or other arrangement for the performance of mechanical work. It is adequately ventilated by a current of air passing through an apparatus similar to that of Benedict, described on p. 785. Any work done is measured and can be calculated in terms of heat, since 1 kg. m. = ·0024 Cal. It is thus possible to estimate simultaneously the total heat production of an individual as well as the respiratory exchange, including both carbon dioxide output and oxygen intake. The general principle of the calorimeter is shown in the diagram (Fig. 526). The calorimeter is also supplied with bed, table, chair, etc., and food can be introduced through a double window so that an experiment may be continued over several days on one and the same individual.

INDIRECT CALORIMETRY. From the equation

$$\underset{(180\text{ g.})}{C_6H_{12}O_6} + \underset{(134\text{ l.})}{6O_2} = \underset{(134\text{ l.})}{6CO_2} + 6H_2O$$

and from the observation that 1 g. of glucose on combustion yield 3·7 Calories of heat, we can deduce that when carbohydrate is being burnt in the body (*i.e.* when the R.Q. = 1·0), the utilisation of each litre of oxygen generates about 5·05 Calories of heat. Similar calculations can be made for the combustion of proteins (4·6) or fats (4·74), so that in all cases the calorific value of oxygen is in the neighbourhood of 5 Calories per l., the exact value of the non-protein part varying according to the respiratory quotient, as shown in the following Table (Cathcart and Cuthbertson) (*v.* also p. 805).

Non-protein R.Q.	Cal. per litre oxygen.
1.00	5·047
0·90	4·936
0·80	4·827
0·718	4·735

These figures have been obtained by actual comparison between respiratory exchange and calorimeter experiments on man. It is therefore now quite easy, when the oxygen usage and R.Q. are known, to calculate the total energy liberation of the body.

Careful experiments have shown that the sum of the energies put out by the body is equal to the sum of the energy obtained by the oxidation of the tissues and of the foodstuffs in the body during the same time. Here are the figures from one such experiment by Atwater on man :

	(a) Heat of combustion of food eaten.	(b) Heat of combustion of fæces.	(c) Heat of combustion of urine.	(d) Estimated heat of combustion of protein gained (+) or lost (−).	(e) Estimated heat of combustion of fat gained (+) or lost (−).	(f) Estimated energy of material oxidised in the body; a − (b + c + d + e)	(g) Heat determined.	(h) Heat determined greater (+) or less (−) than estimated; g − f.	(i) Heat determined greater (+) or less (−) than estimated: h ÷ f.
Date.	Cals.	Cals.	Cals.	Cals.	Cals.	Cals.	Cals.	Cals.	%
Dec. 9–10	2519	110	142	−85	+ 3	2349	2414	+65	+2·8
10–11	2519	110	133	−25	−44	2345	2386	+41	+1·7
11–12	2519	110	132	−21	−93	2391	2413	+22	+0·9
12–13	2519	110	133	−14	−55	2345	2375	+30	+1·3
Total 4 days	10,076	440	540	−145	−189	9430	9588	+158	
Average one day	2519	110	135	−36	−47	2357	2397	+40	+1·7

Basal Metabolism. During fasting the energy output of the body is determined (*a*) by the amount necessary to keep the body alive—*i.e.* to maintain its warmth and to furnish the energy for respiratory movements, contractions of the heart, etc. ; (*b*) the energy necessary to carry out any external work that is performed.

The energy requirements under (*a*) represent the " Basal Metabolic Rate " (B.M.R.) of the body. It may be defined as the energy expenditure of the body in the resting state, a few hours after the absorption of the last meal and under comfortable conditions of temperature, etc. The basal metabolism, of course, varies with the size of the individual ; but if a number of different animals are compared, the smaller the animal, the greater their metabolism per kilo. is found to be.

This is due to the fact that the chief expenditure of energy is devoted to maintaining the temperature of the body, and the smaller the animal the larger is its surface, and therewith its heat loss, relatively to its weight. The basal metabolism, therefore, is more nearly a function of the surface of the body. The surface area, S, of an animal can be found approximately by the Meeh formula * : $S = k\sqrt[3]{W^2}$, where W is the weight and k a constant for the species, *e.g.* for cat, 0·099 ; rabbit, 0·125 ; rat, 0·091 ; dog, 0·107 ; man, 0·123. (Probably 0·11 is more nearly correct for man.) A comparison of the metabolism in different warm-blooded animals shows that it is of the same order in all cases, viz., a daily expenditure of about 700 to 1,000 Calories per square metre body surface, in order to maintain its temperature and carry out such motor processes as are essential to life.

---

* Meeh. *Zeit. f. Biol.*, 1879, **15**, 425. For dogs, Cowgill and Drabkin's formula (modified) is better : A (in m²) = $0{\cdot}2864\ W^{0.367}\ L$, where W = weight in kg. and L = body length in m. from nose to anus (*Am. J. Physiol.*, 1927, **81**, 36).

The same thing applies to man. If we desire to predict the normal basal metabolism of any given individual we must find his surface. This depends

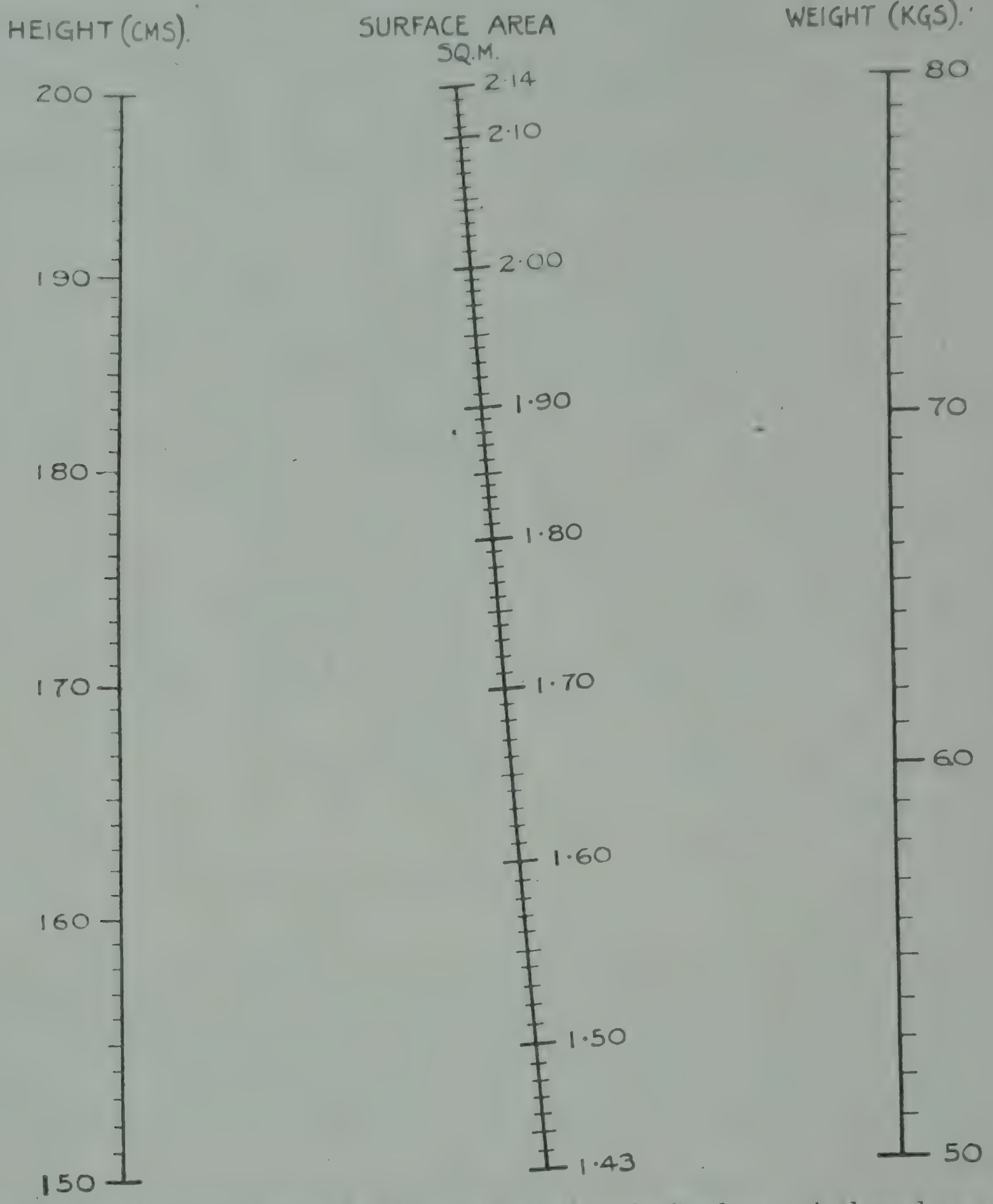


FIG. 527. Alignment chart for Calculation of Body Surface. A thread or ruler, placed so as to lie on the points for weight and height, cuts the line for surface at the correct value. (W. A. M. SMART.)

on height as well as on weight. If we know the height and weight, the surfac can be calculated by the following formula (Du Bois *):

$$S = \cdot 007184 \times W^{0.425} \times H^{0.725}.$$

where S is the surface in square metres, W the weight in kilogrammes, H th height in centimetres ; or, more conveniently, by the use of a nomogram suc as is shown in Fig. 527.

In young adult white men the basal metabolism is about 40 *Calories pe square metre per hour*. It is about 2·5 Calories less for women. From the age o forty onwards there is a reduction, the formula † being 39·138 — 0·0678 × age Cal./sq. m./hr., *e.g.* at fifty it would be about 36 Cal./sq. m./hr.

In the new-born infant, however, it is low (25 Calories per square metr per hour), and is even lower in premature infants. It rises rapidly in the firs year of life, and at five years is about 52 Calories per square metre per hour

---

* DU BOIS and DU BOIS. *Arch. Int. Med.*, 1916, **17**, 863, 887.
† LEWIS. *Am. J. Physiol.*, 1938, **121**, 502.

There are racial variations, due partly to varying diet and in part to different extents of muscular relaxation.

It is usual to determine basal metabolism indirectly by estimating the oxygen usage of the resting, fasting individual, and calculating from the calorific value of oxygen at the observed respiratory quotient. The observed value can then be compared with the value predicted from the body surface. The results are usually given in percentages of the normal predicted basal metabolism. In certain diseases the normal is departed from, a divergence of 10 per cent. being considered significant. Thus in fevers there is an increase during the febrile state, which may amount to 50 per cent. In diseases associated with over-action of the thyroid there is also an increase to 20 to 50 per cent. above normal, while in the converse condition of subthyroidism there is a reduction of 10 to 30 per cent.

Injection of adrenaline, or mental excitement increases the basal metabolism.

Since the cardiac minute volume varies directly with the oxygen usage of the body observations of pulse rate and pulse pressure often give a clue to the B.M.R.* If, for example, we suppose that the stroke volume of the heart (S.V) is proportional to pulse pressure (P.P.) and to surface area (S.A.), and since minute volume = pulse rate (P.R.) × S.V., then the O_2 used, $\propto$ P.R. × S.V. $\propto$ P.R. × P.P. × S.A.

For Reference

W. M. Boothby and Irene Sandiford (1924). Basal Metabolism. *Physiol. Rev.*, **4**, 69.

Carpenter (1924). Tables, factors and formulæ for computing respiratory exchange and biological transformations of energy. (*Carnegie Inst. Washington*, Publication 303A.)

Kleiber (1947). Body Size and Metabolic Rate. *Physiol. Rev.*, **27**, 511.

Newburgh (1944). Obesity. *Physiol. Rev.*, **24**, 18.

F. B. Talbot (1925). Basal Metabolism of Children. *Physiol. Rev.*, **5**, 477.

Terroine et Zunz. Le Métabolisme de Base. (*Press. Univ. de France.*)

3. INFLUENCE OF STATE OF NUTRITION ON METABOLISM

The basal metabolism depends on the state of nutrition of the animal. Thus in fattening an animal, the basal metabolism increases steadily as the animal gets fatter. It requires much more food to keep a fat animal fat, than to maintain it in a lean condition. The figure 40 Cal./sq. m./hr. applies to a well-nourished individual. If the food supplied to the individual be diminished, he can go on working as before, obtaining his energy at the expense of the fat and muscular tissues of his own body. But when, under this régime, the weight has fallen about 12 per cent. it is found that the basal metabolism is also largely reduced. During the War of 1914–18, Zuntz and Loewi † found their own basal metabolisms lowered by 15 and 12 per cent. In a squad of men investigated by Benedict,‡ the food normally taken ranged from 3,200 to 3,600 Calories per day. After these men had by restriction of diet to 1,400 Calories suffered a loss of 12 per cent. in their body weight, the basal metabolism was 18 per cent. below normal. They could maintain themselves at this reduced weight on a diet of 1,950 Calories and could carry out the same work as before on a diet of 2,300 Calories. With further deprivation, when the body weight has been reduced by 20 per cent., there may be a return of the basal metabolic rate up to or even above the normal level.

How great a reduction of calorie intake over long periods is compatible with life is uncertain. In World War I, during four years, the average daily Calorie intake of the Belgians fell from 2,850 to 2,250 ; in World War II it had nominally fallen to 950 in 1943 § ; in France, in 1943, it was about 1,000 ; in German prison and extermination camps it must have been much less. During the siege of Madrid in 1936–39 also, the caloric intake fell to about 1,000, with 66 g. protein, of which animal protein was about 5 g. daily.*

---

* Read and Barnett. *Arch. Int. Med.*, 1936, **57**, 521.

† Zuntz and Loewi. *Biochem. Ztschr.*, 1918, **90**, 244.

‡ Benedict *et al.* *Carnegie Inst., Publ. No.* 280, 1919.

§ Bigwood. *B.M.J.*, May 22nd, 1943, p. 641.

Starvation. During starvation the whole energy expenditure is derived from the animal's own tissues. Even if all the tissue components could be used up, the store of available carbohydrate, chiefly as glycogen, would only supply one day's need, whereas fat and protein reserves are much greater, as shown below :—

Storage of Energy in the Adult Human Body (Young †)

Material.	Weight of Material ingested Daily as Food (g.).	Weight of Material in Adult Human Body (g.).	Number of Days' Supply stored in the Adult Body (approx.).
Carbohydrate . .	400	400	1
Fat	100	12,600	126
Protein . . .	100	13,100	131
Salts	27	3,000	111

According to the evidence of fasting men, this is not painful, provided water is available. During the first day or two there is a craving for food at meal-times. This often passes off, and later even the desire for food may be absent. The essential characteristic of inanition is an ever-increasing weakness, accompanied by a strong disinclination to undertake any mental or physical exertion. The time is passed in a state of semi-stupor. The pulse rate and the body temperature remain nearly normal until a few days before death, which is ushered in by an increase in the somnolent condition of the animal and by a gradual slowing of respiration and fall of temperature. The urine is reduced with diminution in the output of urea and in the amount of water consumed. Some fæces are formed, and may be voided during or at the close of the starvation period. Especially instructive in this connection are the cases of the professional "fasting men," several of whom have been subjected to metabolic investigation. In the case of Succi, who fasted for thirty days, considerable muscular exertion was, by an effort of will, undertaken on the twelfth and on the twenty-third day of starvation without any appreciable ill-effects. In Succi fæces varied from 9·5 to 22 g. a day and contained from 0·3 to 1·0 g. nitrogen per day.

During the whole of the period there is a steady loss of body weight, *e.g.* during the first ten days, the daily loss is between 1 and 1·5 per cent. of the original total weight. This loss of weight does not affect all parts of the body alike. It might be imagined that those organs which are most active should show also the greatest loss of weight. The very reverse of this is the case, as will be seen from the Table on p. 797.

Those organs which are most necessary for the maintenance of life, the brain, the heart, the respiratory muscles, undergo very little loss of weight. Of the other tissues the fat, which is a mere reserve to provide for such contingencies, is drawn upon first, and during starvation 97 per cent. of the total fat of the body may be consumed. The nitrogen needs of the body seem to be supplied chiefly at the expense of the muscles and glands. Hence it is easy to understand the muscular inactivity. We may suppose that during starvation, all tissues undergo a process of slow autolysis, giving up the products of this process to the blood and that the active tissues take up the material set free from all the other cells of the body and so maintain their weight at the expense of all other parts.

The animal organism, in starvation, deals with the resources of its tissues

---

* Marrack. *Brit. J. Industr. Med.*, 1944, **1**, 114.

† Young. *Sci. Prog.*, 1948, No. **141**, 13.

PERCENTAGE LOSS OF WEIGHT OF DIFFERENT ORGANS AND TISSUES DURING STARVATION (VOIT)

Tissue.	Percentage loss of weight of fresh tissue.	Percentage loss of dry tissue.
Fat	97	—
Spleen	67	63
Liver	54	57
Testes	40	—
Muscles	31	30
Blood	27	18
Kidneys	26	21
Skin and hairs	21	—
Intestine	18	—
Lungs	18	19
Pancreas	17	—
Heart	3	—
Brain and spinal cord	3	0

with the utmost possible economy. The total metabolism sinks rapidly during the first two days of starvation, and then more slowly, but the metabolism per kilo. body weight till within a day or two before death is not changed so much. This is shown in the following Table of the output of energy in man during a five days' period of starvation (Tigerstedt):

METABOLISM DURING STARVATION (MAN)

Day of experiment.	Nitrogen output.	Fat oxidised.	Total Calories.	Calories per kilo. body weight.
1	12·16	204·8	2231	33·3
2	12·85	190·3	2112	32·1
3	13·62	179·9	2032	31·3
4	13·67	176·4	2003	31·3
5	11·44	180·0	1979	31·4

THE METABOLISM OF CARBOHYDRATE, FAT, AND PROTEIN DURING STARVATION. During the first phase of inanition the carbohydrate stores in the body rapidly diminish. This stage lasts for one or two days at most, after which energy is derived mainly (80–90 per cent.) from fat and slightly (10–20 per cent.) from protein. In the final stages, as fat becomes scarce, the energy is derived to an increasing extent from oxidation of protein. After prolonged undernutrition there is a lowering of the plasma albumin.

The relative parts played by protein, carbohydrate and fat respectively in the chemical exchanges of a starving animal may be determined in the following way: The amount of protein utilised is given by estimating the total nitrogen of the excreta. The oxygen intake and carbon dioxide output by the lungs must also be determined. Now we know that for each gramme of urinary nitrogen there has been used 5·923 l. of oxygen to oxidise the protein, and 4·754 l. of CO_2 have been produced thereby. The amounts of oxygen and CO_2 which represent protein combustion can thus be calculated, and these quantities are now deducted from the total exchanges, and the remainders expressed as a quotient, $\frac{CO_2}{O_2}$ and called the *non-protein respiratory quotient*. By a simple calculation, the amounts of carbohydrate and fat burned can then be found; more usually a table is used, such as the following:

VALUE OF 1 L. OF OXYGEN IN NON-PROTEIN RESPIRATORY QUOTIENT
(After MCCLENDON)*
One litre of oxygen is equivalent to

Non-protein. R.Q.	Grams.		Calories.
	Carbohydrate.	Fat.	
1·00	1·232	0·000	5·047
0·95	1·010	0·091	4·985
0·90	0·793	0·180	4·924
0·85	0·580	0·267	4·862
0·80	0·375	0·350	4·801
0·75	0·173	0·433	4·739
0·707	0·000	0·502	4·686

For example, suppose that a man in twenty-four hours excretes 10 g. N, uses 426 l. of oxygen, and produces 322·5 l. of CO_2.

Then the oxygen and CO_2 corresponding to the nitrogen output would be 59·23 l. and 47·54 l. respectively. Deducting these from the respiratory exchanges we have a non-protein exchange of:

$$322{\cdot}5 - 47{\cdot}5 = 275 \text{ l. } CO_2 \text{ and } 426 - 59 = 367 \text{ l. } O_2$$

with a non-protein R.Q. of $\frac{275}{367} = 0{\cdot}75$

and for this each l. of oxygen = 0·173 g. carbohydrate and 0·433 g. fat. Hence the material used is:

10 × 6·25 g. = 62·5 g. protein,	which = 62·5 × 4·1 =	256 Cal.
367 × 0·173 g. = 63·5 g. carbohydrate,	which = 63·5 × 4·1 =	260 Cal.
367 × 0·433 g. = 159 g. fat,	which = 159 × 9·3 =	1,480 Cal.
	Total Calories =	1,996

If, as we suppose, the subject is fasting, the carbohydrate used would be glycogen.

A check on this calculation is furnished if the individual can be placed at the same time in a calorimeter.

The influence of glycogen breakdown on that of protein during the first and second days of fasting is shown in the experiments given below:

	First day.			Second day.		
	Glycogen metabolised.		N. eliminated.	Glycogen metabolised.		N. eliminated.
	Total.	Per kilo.		Total.	Per kilo.	
S.A.B.	181·6	3·15	5·84	29·7	0·52	11·04
S.A.B.	135·3	2·31	10·29	18·1	0·31	11·97
S.A.B.	64·9	1·09	12·24	23·1	0·39	12·45
H.C.K.	165·6	2·33	9·39	44·7	0·64	14·36
H.R.D.	32·8	0·59	13·25	41·6	0·76	13·53

The relative part taken in the production of the total energy by fats and proteins respectively may vary according to the amount of fat available in the body. If the subject previous to the experiment has been receiving a diet rich in protein, the excretion of nitrogen and urea in the urine diminishes rapidly during the first days of starvation. During the first two days, therefore, a considerable proportion of the necessary energy is obtained at the expense of protein. Between the third and fifth day the nitrogenous excretion reaches a minimum, at which point it remains approximately constant until a day or two before death. If the previous diet was very poor in protein, the excretion of nitrogen may be low throughout the whole course of the experiment. These facts are illustrated by the curves in Fig. 528, which show the output of urea in three experiments on a dog on high, medium and low previous diets of protein. These experiments indicate that it is not possible, by giving extra protein in the diet, to provide any large store of body protein. The potassium excretion in starvation runs parallel to that of nitrogen, and corresponds to the potassium content of the lost muscular tissue.*

---

* MCCLENDON and MEDES. "Physical Chemistry in Biology and Medicine."

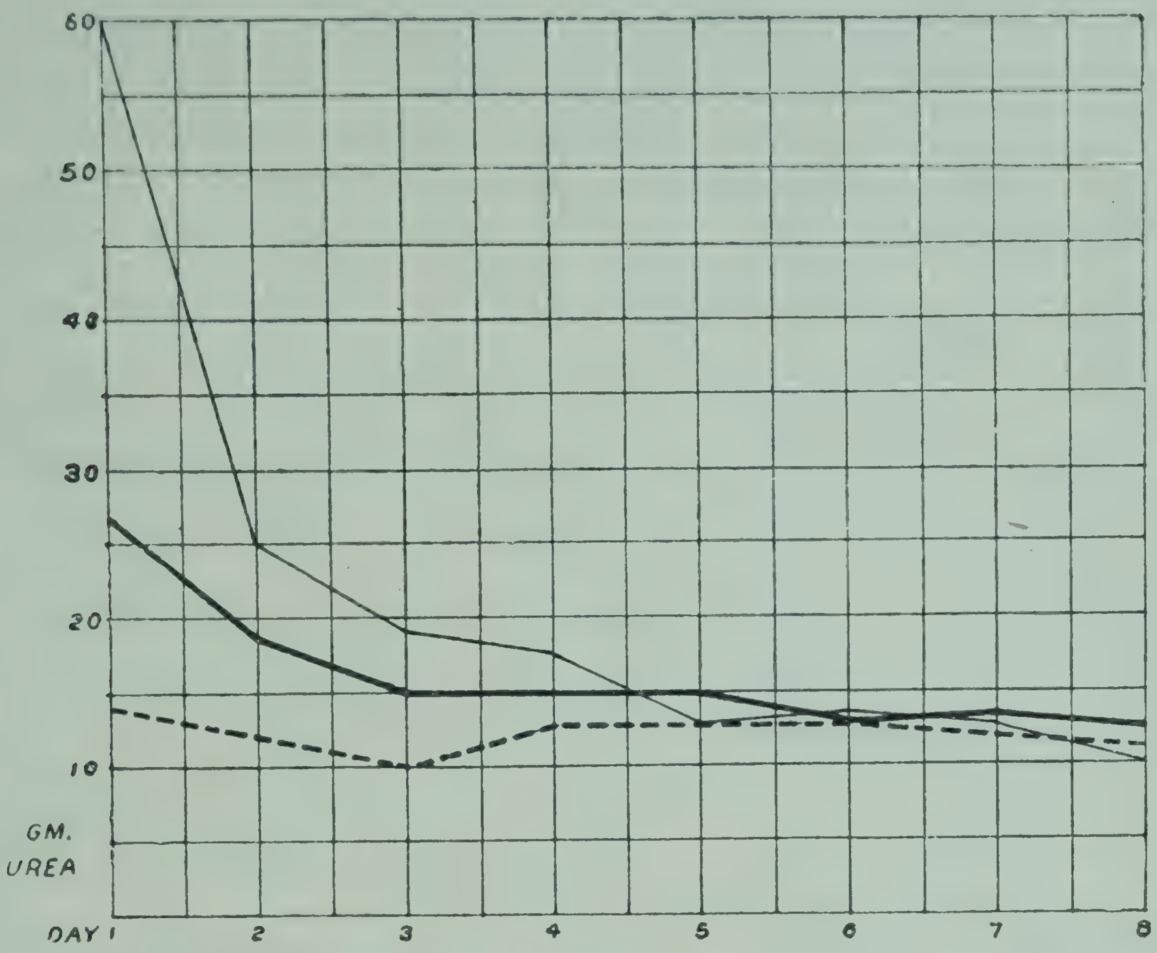


Fig. 528. Three Experiments on the Output of Urea during Starvation (dog). (Tigerstedt, after Voit.)

In (1) (thin line), the dog received 2,500 grammes meat per day before the experiment; in (2) (thick line), the diet was 1,500 grammes meat; and in the third experiment (dotted line), the meat was reduced to a minimum.

In prolonged fasting the nitrogen output falls gradually to a low level as is shown below.

Daily Nitrogen Excretion of Succi in Starvation

Day.	N.	Day.	N.	Day.	N.
1	17·0	8	9·74	15	5·05
2	11·2	9	10·05	16	4·32
3	10·55	10	7·12	17	5·4
4	10·8	11	6·23	18	3·6
5	11·19	12	6·84	19	5·7
6	11·01	13	5·14	20	3·3
7	8·79	14	4·66	21	2·82

Often just before death there is a rise of protein metabolism, synchronous with a practically complete disappearance of fat from the body. The animal now has to supply all its requirements at the expense of protein, which accounts for the increased excretion of nitrogen. This is shown in the following experiment of Rubner on a rabbit:

Days.	Average daily output of nitrogen.	Average amount of fat oxidised daily.
1–3	1·67 g.	10·3 g.
4–5	1·46 ,,	10·3 ,,
6–8	3·21 ,,	2·4 ,,

We see therefore that during starvation, apart from the first day or two, the animal derives the main portion of its necessary energy from the combustion of fats, provided that there is a sufficient store of these substances in the body. Since protein comes from the working tissues of the body, these are spared so far as possible, and it is only when the stored fat is used up that any large call is made on the tissue protein.

---

* Gamble, Ross and Tisdall. *J. Biol. Chem.*, 1923, **57**, 633.

It is probable that part at least of the protein that becomes available in food deprivation comes from the liver. Some also comes from blood plasma, which shows a fall in serum albumin. Hunger œdema, which is especially associated with chronic protein shortage, appears to be a result of this.

In starvation the ability to metabolise carbohydrate is depressed, so that if glucose is administered, there is an abnormally high blood sugar, and glucose is excreted in the urine—so-called hunger diabetes.

For Reference

Chambers (1938). Undernutrition and Carbohydrate Metabolism. *Physiol. Rev.*, **18**, 248.
G. Lusk (1921). The physiological Effect of Under-Nutrition. *Physiol. Rev.*, **1**, 523.
S. Morgulis (1923). Fasting and Under-Nutrition. (Dutton, New York.)

4. EFFECT OF THE FOODSTUFFS ON METABOLISM

Food not only supplies energy, but also serves to replace the wear and tear of the body. So far as the latter function is concerned the proteins of the food take a position apart. Every working cell of the body consists largely of proteins ; to repair the wear and tear of the machine of the body, therefore, proteins are absolutely necessary. To supply the energy requirements any or all of the three classes of foodstuffs can be utilised.

The Influence of Proteins. A starving man is putting out energy and diminishes in weight as a result. In the example cited on p. 798 the man was losing about 13 g. of nitrogen a day and was putting out about 2,000 Calories, derived partly from the combustion of protein and partly from the oxidation of the fat. It would manifestly be useless to try to stay the loss from the body by giving merely 13 g. of nitrogen in the food in the form of protein, since this would give him only 320 Calories towards his basal requirements of 2,000. We should find in fact that the effect of such an administration would be practically to double the output of nitrogen from the body and to increase slightly the total output of energy. In the dog, it would be possible to stop the loss of body weight by giving an amount of protein equal to five times the loss during starvation. In man, such an amount would be too great for his assimilating powers and it would be necessary to give fats and carbohydrates in addition to the protein.

Nitrogenous Equilibrium. A subject is said to be in a state of nitrogenous equilibrium when the amount of nitrogen excreted during a given period is equal to that ingested. The protein taken in with the food has a twofold function to perform. One of the functions is the replacement of tissue. When, however, we are feeding an animal on a pure protein diet, by far the larger portion of the food is utilised for meeting the energy requirements of the body. In this function protein presents no apparent advantages over the other two classes of foodstuffs. Its value to the animal, represented numerically by its physiological heat value as 4·1, is equivalent to the value of the same weight of carbohydrate and is far inferior to that of fat. If, instead of giving to the starving animal a pure protein diet, we administer a mixed diet containing fat or carbohydrate, or both substances, sufficient to meet the normal energy requirements of the body, we can restrict the utilisation of protein more nearly to the replacement of tissue waste in the body, and are therefore able to attain nitrogenous equilibrium with a much smaller proportion of protein than is possible when this substance furnishes the whole diet. In omnivora, such as man, it is easy to attain nitrogenous equilibrium on a mixed diet with a smaller nitrogen turnover than is found during starvation. In the experiment given on p. 798 the average nitrogen output during fasting was about 13 g., and in prolonged starvation it may fall to 5 g. per day or less.

Chittenden * found that in man apparently normal nutrition may be maintained on a mixed diet containing only 7 g. of nitrogen daily. If the amount of fat and carbohydrate be very largely increased, it is possible to maintain nitrogenous equilibrium on even smaller quantities of protein. Thus Kon and Klein † kept a man and a woman, each of about 63 kg. body weight, for 167 days in good health and nitrogenous equilibrium on an average daily nitrogen intake of 5·7 and 3·8 g. respectively.

Most of the experiments on the influence of variations in the quantity of protein food have been made on carnivora, such as the dog and cat. Within very wide limits, the output of nitrogen is proportional to the intake. This is shown in the Tables by Voit ‡ given below, representing two experiments on dogs.

In Experiment I. the animal had been fed for some days before with 500 g. of meat (17 g. N) per diem. The fact that he was excreting 18·6 g. nitrogen daily shows that he was not yet in a condition of nitrogenous equilibrium. Each day he was using up 1·6 g. of his body nitrogen, in addition to the 17 g. supplied in the food. On increasing his food threefold to 1,500 g. (= 51 g. N) the nitrogenous output was considerably increased, but now some nitrogen was retained in the body; the output of N continued to rise, and the amount retained to fall, until on the eighth day of the experiment nitrogenous equilibrium was reached. During the six days intervening 26·4 g. N had been retained in the body, *i.e.* there had been a retention of protein, probably in the form of increased muscular substance.

EXPERIMENT I.				EXPERIMENT II.			
Day.	Daily N. intake. g.	Daily N. output. g.	N-balance. g.	Day.	Daily N. intake. g.	Daily N. output. g.	N-balance. g.
1	17·0	18·6	− 1·6	1	51·0	51·0	0
2	51·0	41·6	+ 9·4	2	34·0	39·2	− 5·2
3	51·0	44·5	+ 6·5	3	34·0	36·9	− 2·9
4	51·0	47·3	+ 3·7	4	34·0	37·0	− 3·0
5	51·0	47·9	+ 3·1	5	34·0	36·7	− 2·7
6	51·0	49·0	+ 2·0	6	34·0	34·9	− 0·9
7	51·0	49·3	+ 1·7				
8	51·0	51·0	0				
		Total	+ 26·4			Total	− 14·7

In Experiment II. diminution of the food (from 1,500 g. to 1,000 g.) was followed by a loss of N. from the body, the output being greater than the income. The excess, however, was rapidly diminishing and equilibrium had been practically attained on the last day of the experiment. During this time the animal had excreted 14·7 g. more nitrogen than it had received in its food. Many such experiments have agreed in showing that in carnivora an appreciable storage of nitrogen can take place in the body, probably by a growth of muscle substance. However much protein is given, the limit is finally arrived at where no further laying on of protein tissues of the body is possible, and the animal then enters into a state of nitrogenous equilibrium. This equivalence of income and output signifies that the *total energy output* of the body rises and falls with the quantity of protein in the food. This is shown in the Table on p. 802 by Pettenkofer and Voit,§ in which the figures were recalculated by Pflüger.

We see, therefore, that *carnivorous animals* can satisfy their total energy requirements at the expense of protein. When the protein income is in excess of their requirements, a small amount is laid on, probably as increased muscular tissue. The most marked effect, however, is an increased metabolism, which rises in proportion to the nitrogenous income. The limit to this increase is set by the powers of the alimentary canal to digest the protein. The rise in metabolism consequent on protein food is very rapid and affects the gaseous exchanges as well as the output of nitrogen. Magnus Levy and Falk found that a large protein meal might increase the respiratory exchange 40 per cent., during seven hours.

---

* CHITTENDEN. "Physiological Economy in Nutrition" (New York). 1905.

† KON and KLEIN. *Bioch. Journ.*, 1928, **22**, 258.

‡ VOIT. *Ztschr. f. Biol.*, 1867, **3**, 79 (recalculated by Tigerstedt).

§ PETTENKOFER and VOIT. *Ztschr. f. Biol.*, 1869, **5**, 370.

Period.	Nitrogen in food.	Nitrogen output.	Fat gain or loss.	Total Calories.
1	0	5·6	−98	1072
2	17	20·4	−59	1094
3	34	36·7	−38	1328
4	51	51·2	−22	1554
5	61	59·7	−30	1851
6	68	69·5	+22	1600
7	85	85·4	+12	2128

The nitrogenous output also rises immediately after a protein meal, so that 50 per cent. of the nitrogen of the ingesta may appear in the urine within seven hours after the meal.

These results cannot be strictly applied to man, in whom it is impossible to supply all the energy requirements of the body on a pure protein diet. Even if a man eats as much meat as he can, he will be unable to obtain sufficient energy for his daily needs. Whereas the average daily requirements of a man amount to about 3,000 Calories, 1 lb. of lean meat (*i.e.* entirely devoid of fat) would yield only about 400 Calories, and even if he took 4 lb. of meat daily, an amount which is impossible for most individuals, he would be obtaining only about 1,600 Calories. The cures for obesity, in which a large protein diet plays an important part, are in all cases practically equivalent to a state of semi-starvation.

Many experiments have been made on the influence of variations in the quantity of protein in a mixed diet. Within wide limits the output of nitrogen is equal to the intake. A normal adult man seems able to store little, if any, protein. The only way in which protein can be laid on in the body is by furnishing a physiological stimulus to the growth of muscle, *i.e.* by constant exercise. The conditions, however, are different when dealing with an individual in whom, from some cause or other, the muscular tissues are not at their full development. Thus in growing individuals a certain amount of the protein of the food is always retained and laid on as tissue protein. Weight for weight, a boy of sixteen to eighteen requires about 50 per cent. more protein than an adult man. In convalescence after great wasting of the muscles has taken place, feeding with large amounts of protein has been found to give rise to a considerable retention of protein in the body. This process goes on only until the muscles have attained their normal condition of development.

Specific Dynamic Action. Protein food has in man, as in animals, a specific stimulating effect on the metabolism, so that the respiratory exchange is largely increased as a result of a heavy protein meal. The rise is about 30 per cent. of the basal value ; *i.e.* if for each 100 Calories basal metabolism we give protein equal to 100 Calories, then the metabolism, instead of remaining at 100, rises to 130 Calories. This effect was named by Rubner the "*specific dynamic action*" of protein.

If either fats or carbohydrates be given to a starving animal, a certain sparing of the fat of the body takes place, but this effect is accompanied by a distinct specific dynamic action, viz. from a basal value of 100 Calories the metabolism is increased to 106 Calories by 100 Calories glucose, or to 104 by 100 Calories fat. In undernourished individuals the effects are greater than these, and tend to bring the metabolism up to what it would be if the basal level were normal. As regards the protein metabolism, Cathcart found that while administration of fat increases the nitrogen output during starvation, carbohydrate food causes a diminution in the nitrogen output, and thus exercises a marked sparing effect on the proteins of the body.

The stimulating influence of carbohydrate on metabolism, although much less than that observed on the administration of protein, is easily demonstrable after administration of moderate amounts of soluble carbohydrate to a fasting subject. If carbohydrate be given in excess of the daily energy requirements, the greater proportion of it remains in the body, being stored up provisionally as glycogen, but eventually in the form of fat.

If excess calories be given in the form of carbohydrate, there is a sparing of protein, and N. and S. are stored. Conversely, if the protein and carbohydrate portions of a diet are consumed at different times, there is a loss of the order of 2 g. N. daily. It is clear, therefore, that the metabolism of carbohydrate and of protein are linked in some way: it may be that intermediaries of carbohydrate take the place of amino-acids in being oxidised.*

Cause of Specific Dynamic Action. The metabolic rate during the absorption of a meal is always greater than that under otherwise comparable conditions in the fasting subject. We have seen that the specific dynamic action of the various foodstuffs results in the augmentation of the basal, *i.e.* fasting, rate of metabolism by the following percentage amounts :—

Protein	30 per cent.
Carbohydrate . . .	6 ,,
Fat	4 ,,

The augmentation is not so constant as these figures suggest, however. With protein it varies according to the amount given, first decreasing, then reaching a level and finally increasing as the total caloric contribution from protein is raised.

The explanation of specific dynamic action is not entirely clear and the effect may be due to various causes, but in general it may be said to be a consequence of the energy expenditure involved in processes of intermediary metabolism. Since protein has the most striking effect it has received most attention. The effect is also produced by some of the amino-acids, *e.g.* glycine, alanine, tyrosine and phenyl-alanine and to a smaller extent by leucine and glutamic acid : it has its seat mainly in the liver, and seems largely to represent heat set free in the conversion of the amino-acids to urea. Two stages of this conversion involve energy expenditure, first the liberation of ammonia by an oxidative change, and second, the coupled oxidative reactions associated with conversion of the ammonia to urea. If the amino-acids are not broken down, but are built into the tissues, as in rapidly growing young animals, no specific dynamic action is seen. It is abolished by removal of the liver, because it is there that the changes occur.

The specific dynamic action of protein results only in the production of heat, and since in cold surroundings a large proportion of the food must always be applied to maintain the body temperature, the excess heat involved in the consumption of protein, and to a less extent of other foods, can be, and is, utilised. Hence it comes about that in order to exhibit this specific dynamic action, the animal must have a temperature above 33° Centigrade. Below this temperature the specific dynamic action of the foodstuffs becomes less and less apparent, and is finally merged in the heat production necessary, whether food is taken or not, to keep up the temperature of the body.

The effect of carbohydrate probably represents the energy expenditure needed to convert it into glycogen, or into fat. The latter involves the greater energy expenditure : if conversion to fat is encouraged by giving vitamin B_1 with the glucose, the S.D.A. is increased.† It is not abolished by liver removal.

The effect of fat is unexplained, and is supposed to be due to increased fat combustion.

For Reference

Borsook (1936). The Specific Dynamic Action of Protein and Amino-acids in Animals. *Biol. Rev.*, **11**, 147.
G. Lusk (1931). The Specific Dynamic Action. *Journ. Nutrition*, **3**, 519.
Schaeffer and Le Breton (1938). "L'Action dynamique spécifique des Protids." (Paris. Hermann.)
Wilhelmj (1935). The Specific Dynamic Action of Food. *Physiol. Rev.*, **15**, 202.

---

* Cuthbertson and Munro. *Bioch. J.*, 1939, **33**, 128.

† Ring. *Amer. J. Physiol.*, 1943, **138**, 488.

5. THE EFFECT OF MUSCULAR WORK ON METABOLISM

The energy for muscular work is immediately derived from the oxidation of glycogen, but ultimately of the foodstuffs. The following Table (Benedict and Cathcart) represents the respiratory exchanges in a trained muscular subject during complete rest and while riding a stationary bicycle. The "severe work" in this Table was carried to exhaustion, so that the respiratory exchanges represent the maximal for that type of exercise.

Condition.	c.c. CO_2 eliminated per minute.	c.c. Oxygen absorbed per minute.	Respiratory quotient.	Pulse rate.	Respiration. Rate.
Lying . .	200	243	0·83	56	20
Moderate work .	1720	1834	0·94	150	32
Severe work .	3227	3265	0·98	166	38

We can determine the relation between work done and energy expended by either direct or indirect colorimetry, as explained on p. 792. In the Table below are given the results of an experiment by Atwater by the direct method. The work was done by the man on a bicycle ergometer.

In these experiments, the difference in the mean expenditure of energy on the work days and on the rest days (with food) was 2,723 Calories. This *excess metabolism* over resting requirements took place in order that the body might perform work equal to 546 Calories. The fraction $\frac{\text{work}}{\text{energy of excess metabolism}}$ is spoken of as the "mechanical efficiency" of the body. In the above case it equals one-fifth, and the body is said to have a 20 per cent. efficiency. Of course a certain proportion of this excess of energy used over work done is accounted for by the increase in the work which must be performed by the respiratory muscles and heart. Even if we neglect these factors altogether, the efficiency of the body as a machine corresponds to between 16 and 25 per cent., an efficiency which exceeds that of the best of our steam engines and is only equalled by certain internal combustion engines.

ENERGY PER DAY.

Nature of experiment.	Heat eliminated.			External work in Calories.	Total in Calories.
	By radiation and conduction.	In urine and fæces.	In water vaporised from lungs and skin.		
Rest with food (average of four days) . . .	1850	26	521	—	2397
Rest, fasting (four experiments) (average of five days)	1605	21	561	—	2187
Work (fourteen experiments) (average of forty-six days)	3802	29	743	546	5120

We have already seen in Book II that the results of metabolism experiments during work may throw light on the nature of the constituents of the body which undergo oxidation to furnish the energy of muscular work. In the experiment already quoted, the oxidation of carbon was increased more than threefold during a six-hour period of work. This carbon might form part of the molecule of protein, fat or carbohydrate ; but it has long been known that muscular exercise causes no considerable increase in the nitrogen excretion.* The above experiments illustrate this fact. The basal output of nitrogen per day corresponded to the destruction of 82 g. of protein. During rest and with food, the average daily consumption of protein was 98·8 g. During a work day, in which the individual received protein and a somewhat insufficient quantity of carbohydrates and fats, the consumption of protein was 109·4 g. Thus there was a three- to four-fold increase of the carbon metabolism of the body, but only a 10 per cent. increase in the protein metabolism.

The study of the respiratory quotient over periods of several hours might be expected to give us some idea of the nature of the material which is furnishing by its oxidation the energy for muscular work. If the whole of the energy requirements were furnished by the oxidation of carbohydrates, $C_6H_{13}O_6 + 6O_2 = 6H_2O + 6CO_2$, the *respiratory quotient* of the animal, namely $\frac{CO_2 \text{ expired}}{O_2 \text{ inspired}}$, would be equal to unity. From the formula of a fat it is evident that it will require a relatively larger amount of oxygen to oxidise it than is necessary in the case of the carbohydrate. The respiratory quotients for protein, fats and carbohydrates are given in the following Table :

Material.	Respiratory quotient $\frac{CO_2}{O_2}$ (volumes).
Carbohydrate	1·0
Protein	0·81
Human fat	0·718

Since the carbon dioxide excreted during exercise may result from the oxidation of protein, fat or carbohydrate, we cannot exactly deduce the total energy set free in the body by an estimation of the O_2 intake alone. The respiratory quotient must also be determined, and the output of nitrogen in the urine measured. Then the non-protein R.Q. is calculated as already explained on p. 797, and from this the relative contribution of the various foodstuffs can be deduced by reference to a table such as the one following, which gives the relative oxygen usage for combustion of carbohydrate and fat for a given non-protein respiratory quotient in man † :

Non-protein R.Q.	Carbohydrate.	Fat.	Calories for one litre oxygen.
1·00	100 per cent.	0 per cent.	5·047
0·95	82·3 ,,	17·7 ,,	4·992
0·90	65·5 ,,	35·5 ,,	4·936
0·85	46·8 ,,	53·2 ,,	4·881
0·80	20·1 ,,	70·9 ,,	4·827
0·75	11·4 ,,	88·7 ,,	4·770
0·718	0 ,,	100 ,,	4·735

The number of Calories set free in the body for every litre of oxygen taken in are also given in the above Table.

The respiratory quotient of an animal is a net result determined by the average nature of the substances which are undergoing oxidation in its body. If the performance of muscular work involved special chemical processes, *e.g.* a metabolism of one of the main constituents of the body in preference to either of the others, this sudden change in the quality of the metabolism should show itself in the respiratory quotient. It must be borne in mind that the respiratory quotient gives only the final result of the oxidative changes, and tells us nothing regarding the intermediate stages by which that end result is reached. For instance, if fat were converted to carbohydrate and the carbohydrate all burned, the respiratory quotient would be the same as if the fat had been burned directly.

Prolonged moderate exercise, although attended by a large increase in the CO_2 output and oxygen intake of the body, does not alter the respiratory quotient.* The respiratory quotient over a long period, such as a day, must depend on the material which is combusted, irrespective of the amount of muscular work which is done in the time. Ultimately there is no doubt that the energy for muscular exercise can be derived from any of the three classes of food. Such experiments, however, do not answer the question as to the immediate origin of the energy involved in muscular contraction. When dealing with this we saw that the ultimate source of the energy of contraction was the glycogen of the muscle. We shall see later on that the body has the power of converting carbohydrates into fats ; and proteins, and perhaps fats, into carbohydrates. The respiratory quotient in the course of a day is affected by these intermediate changes as well as by the foodstuffs oxidised.

---

* Fick and Wislicenus. *Vierteljahrs. d. zürch. Naturf. Gesselsch.*, 1865, **10**, 317.

† Cathcart and Cuthbertson. *Journ. Physiol.*, 1931, **72**, 249. The slightly different values from those on p. 792 is due to the fact that the composition here taken for fat is that of human fat.

As the result of a series of experiments Krogh and Lindhard † came to the conclusion that the cost of work in man, when it is performed at the ultimate expense of fat, is 10 per cent. more than when it is performed at the expense of the carbohydrate of the body. This extra expenditure of energy they regard as, so to speak, wasted in the chemical conversion of fats into carbohydrates. This great difference does not arise in the case of proteins, which form carbohydrates in the body with extreme ease. But whereas proteins form, for the most part, the working tissues of the body, fats are stored as a reserve food, so that under normal conditions, in default of an excessive diet of proteins, fats must often be utilised for the carrying out of muscular work. For this to take place, it is probable that they are first converted into carbohydrates and stored in the muscles as glycogen.

It will be noticed that, in contradiction to popular opinion, the proteins do not play any special part in relation to muscular exercise.‡ The special value of protein is for building up the body and for repairing wear and tear. As a source of energy it presents no advantages over carbohydrate and fat ; in fact it may be in some respects inferior to these substances, or at any rate less economical, owing to the specific dynamic action, the whole of which appears in the form of heat. If an animal after a protein meal be made to work, the increased metabolism due to the specific dynamic action of the protein is still observed over and above that due to the performance of work. On the other hand, in training there is some reason for a plentiful supply of protein in the diet. When, however, a man has arrived at the full development of his muscular system, there is no need to attempt to supply the energy for his movements at the expense of protein. In many cases the beneficial effects of meat are due to its content of fat.

For Reference

Cathcart (1925). The Influence of Muscle Work on Protein Metabolism. *Physiol Rev.*, **6,** 225.

G. Lusk (1928). The Science of Nutrition. (Saunders, Philadelphia and London.)

H. B. Richardson (1929). The Respiratory Quotient. *Physiol. Rev.*, **9,** 61.

Steinhaus (1933). Chronic Effects of Exercise. *Physiol. Rev.*, **13,** 103.

6. THE SIGNIFICANCE OF THE FOODSTUFFS

Isodynamic Replacement. A primary requirement of a diet is that it provide an appropriate source of energy. For supplying this energy any of the three classes of foodstuffs may be utilised. Thus carbohydrates and proteins are *isodynamic—i.e.* give rise to the same amount of energy in the body for a unit weight—1 g. of each giving out 4·1 Calories. Within limits, carbohydrate can be replaced by the isodynamic amount of fat, and within narrower limits all three classes are isodynamically interchangeable. This isodynamic equivalence of the foodstuffs is, however, interfered with to a certain extent by the specific dynamic action. On account of this, an average mixed diet, in order to maintain the body in a state of equilibrium, must have a Calorie value about 15 per cent. higher than the basal requirement. Thus, supposing a man while fasting and in bed gives off 1,800 Calories a day, he would under the same circumstances give off about 2,000 Calories if he received a minimum mixed diet just sufficient to maintain his body weight constant.

---

[*] Gemmill. *Amer. J. Phys.*, 1931, **98,** 135.

[†] Krogh and Lindhard. *Bioch. J.*, 1920, **14,** 290.

[‡] Wilson. *J. Physiol.*, 1932, **75,** 67.

Protein. Proteins are distinguished from carbohydrates and fats by the fact that no means are provided for the storage of proteins in the body, except by the growth of muscular and other tissues, and perhaps, to a certain extent, by actual storage of protein in the liver.* Arrived at its maximum state of growth, the body cannot add to its protein store, and therefore all the protein that is ingested and absorbed is at once broken down in the body. If we examine the curve of nitrogenous excretion in the urine, we find that practically all the nitrogen taken in as protein in the food is turned out within twelve to sixteen hours. Thus, if a man is on a mixed diet sufficient for his daily needs, a doubling of the protein in the diet, leaving the other constituents unchanged, will result in a doubling of the nitrogenous excretion. On the other hand, an increase of the fats or carbohydrates will not increase the carbon output of the body, but any excess above the daily needs, if absorbed, is stored up in the body in the form of fat. To increase the protein metabolism, therefore, all that is necessary is to increase the protein intake; to increase the carbohydrate and fat metabolism it is necessary to increase the metabolism as a whole, *e.g.* by muscular exercise or by exposure to cold.

Protein is essential for the repair of the wear and tear of the tissues. We may thus speak of the protein of the food as having a twofold destiny. The amino-acids into which it is resolved circulate in the blood, and there is a continual exchange of amino-acids between the blood and the proteins of the living cells. These proteins are thus maintained by constant intra-molecular repair, or are increased in amount, if growth is in progress, by synthesis anew from the blood amino-acids.

Surplus molecules of amino-acids not used up in these ways are broken down, the nitrogen being converted to urea and excreted, and the remainder of the molecule, after further intermediary changes, being oxidised to CO_2 and water. These aspects of the subject will be dealt with in Chapter XLVI.

Action of the Products of Protein Digestion. Since the proteins of the body differ in their composition from those of the food, it is evident that each protein molecule has to be disintegrated and reconstructed before it can take its place in the body fabric; and it is therefore immaterial, so far as metabolism is concerned, whether we feed the animal with the ordinary food-protein or with the products of its hydrolysis, provided all the essential amino-acids, including threonine, are present; the mixed amino-acids can be given either by mouth or by slow intravenous injection.

The Biological Value of Proteins. Different diets containing the same amounts of protein are not necessarily of equal food value. One reason for this is that the proteins in different foodstuffs show different digestibilities: *e.g.* meat and milk 100 per cent., egg 85, legumes 70–80 per cent., and apple very indigestible. The reason for indigestibility is partly that the proteins are enclosed in indigestible cell walls. Pure proteins show smaller differences. The digestibility of proteins is expressed as the per cent. of consumed N. that does not appear as fæcal N. In consequence of the variation in the composition of the proteins contained in the different tissues of the body, it is necessary that the food shall include certain of the amino-acids in proper quantities. The indispensable amino-acids, ten in number, are called *essential*: the remainder can be synthesised in the body and are hence *non-essential*. Hence it follows that all proteins are not equivalent as regards their capacity for replacing tissue waste or serving for growth. Those in which nearly all the amino-acids are represented and which contain these in proportions approximating the average of those found in the chief tissue proteins are called *first-class proteins*. Most animal proteins belong to this group, though gelatin is an exception; some vegetable proteins such as edestin, glutenin of wheat, or glutelin of maize, are also included.

In the remaining, or second-class proteins, either (1) one or more of the essential amino-acids are absent, or (2) there is a large preponderance of one or more of the amino-acids which are required only in small proportions. As an example of the first type we may instance collagen and gelatin, and zein, the chief protein of maize. As an example of the second type we may instance the protein of wheat flour. Whereas caseinogen contains 22 per cent. of glutamic acid and serum albumin only 8 per cent., gliadin, the chief protein of flour, contains as much as 36 per cent. If wheat flour is the main or sole source of protein to the body, it is evident that a large amount of glutamic acid will be in excess of that required. It is also deficient in lysine.

---

* Foster and Whipple. *Amer. J. Physiol.*, 1922, **58**, 393, 407; Stübel, *Pflüger's Arch.*,1920, **185**, 74.

One of the methods of studying the subject is to feed young animals, *e.g.* rats, on a diet in which the protein lacks certain amino-acids, and to follow the rate of growth. This is reduced if the lacking amino-acid is essential. Gelatin is lacking in cystine, valine, phenylalanine, histidine, methionine, threonine, leucine, isoleucine, tyrosine, tryptophan and hydroxyglutamic acid, and it has long been known that it cannot entirely replace the proteins of the food. Its capacity for yielding energy to the body equals that of any other protein, but it can only replace body protein when the missing amino-acids are also supplied. In the same way zein, which is deficient in lysine and tryptophan, is incapable by itself of supporting life, though this can be accomplished if tryptophan is administered at the same time. Growth does not occur, however, until lysine is also added (Fig. 529).

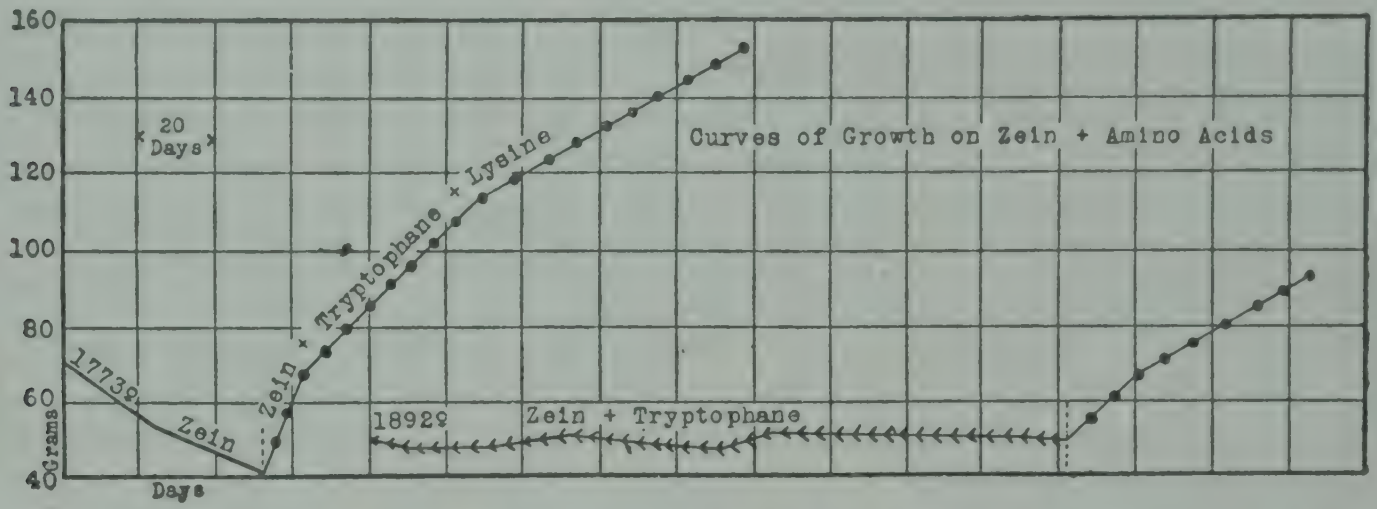


Fig. 529. Growth curves of young rats fed on a diet in which the protein is given as zein. No. 1773 received zein and lost in weight; when tryptophan and lysine were added to the zein it gained weight. No. 1892 received zein and tryptophan for six months; its weight remained stationary, but showed rapid increase when lysine was also given. (Mendel. Courtesy of the *J. Amer. Med. Assoc.*)

Certain proteins are also deficient in the sulphur-containing amino-acid methionine, some in phenylalanine, and some in histidine. All these, together with lysine and tryptophan, are essential, and cannot be made in the body. It is noteworthy that cystine can be replaced by methionine, and tyrosine by phenylalanine; but histidine cannot be replaced by arginine.

Undoubtedly the best method of finding which amino-acids are essential and which are not is to feed young animals on a diet in which protein is replaced by mixtures of pure amino-acids, one of which is omitted. This method has been used with remarkable success by Rose, and threonine was actually discovered in this way, as it was observed that mixtures of all previously known amino-acids were inadequate to promote growth. One of the consequences of lack of essential amino-acids is failure of growth; another is lack of appetite. Lack of valine produces curious rotatory movements in rats. Methionine (or cysteine plus choline, from which methionine can be formed) appears to be essential for the well-being of the liver. Complete lack of methionine causes, in the rat, liver changes ranging from fatal acute yellow atrophy to chronic hepatic cirrhosis.* Liver damage following chloroform anæsthesia, arsphenamine dosage, etc., can be averted by administration of methionine, or of proteins such as casein, which are rich in it.

---

* Himsworth and Glynn. *Lancet*, 1944 i., 457.

The following Table shows the essential and non-essential amino-acids (for rats), according to Rose.*

Essential.	*Non-essential.*
Arginine.	Alanine.
d- or *l*-Histidine.	Aspartic acid.
d-Isoleucine.	Citrulline.
l-Leucine.	Cystine.
d-Lysine.	Glutamic acid.
l-Methionine.	Glycine.
d- or *l*-Phenylalanine.	Hydroxyglutamic acid.
d-Threonine.	Hydroxyproline.
d- or *l*-Tryptophan.	Norleucine.
d-Valine.	Proline.
	Serine.
	Tyrosine.

The original experiments on essential amino-acids were carried out on rats. It has now been shown that most of the amino-acids listed above as essential for rats, are also needed to maintain N. equilibrium in man, so probably there is no great species difference. Lack of valine causes an immediate loss of nitrogen.

These differences in chemical composition between the proteins are of greatest significance when the protein of the food is reduced to the minimal amount required to replace tissue waste, and under such circumstances marked differences are observed between the biological values of proteins. The biological value of a given protein decreases as the amount of that protein is increased.

It is evident that when there is scarcity of protein, animal protein is more economical than vegetable protein; and that if it is necessary to live on a largely vegetable diet this should be mixed, so that all the amino-acids required in the body may be represented in the diet. Thus, the diets at R.A.F. stations in 1941, though only containing about 40 per cent. of the total protein in the form of animal protein, had the high biological value of about 80.†

For the estimation of the biological value of proteins various methods have been used. In one method the nutritive value for growth is expressed in grammes of body weight increase per gramme of protein consumed, when the diet contains that amount of protein at which the maximal growth occurs.‡

The *nutritional biological value* of a protein can be practically assessed by the method of Mitchell.§ It is too involved to be adequately discussed here; its principle is to determine, on rats, what fraction of the absorbed nitrogen escapes excretion in the urine, during the feeding period, *e.g.*

Estimation of N. absorbed and retained.

		Daily. mg. N.
N. in food of diet . . .		=56·9
Fæcal N. on diet . .	=27·6	
„ „ on N.-free diet .	=21·7	
∴ indigestible N. . .	= 5·9	5·9
N. absorbed		=**51·0**
N. wasted		=10·9
N. retained		=**40·1**

Estimation of N. escaping metabolism.

	Daily. mg. N.
Urine N. on test diet . . .	= 48·6
Tissue catabolism, *i.e.* urine N. on N.-free diet	= 37·7
N. wasted in metabolism . .	= **10·9**

$$\text{Biological value} = \frac{40\cdot1}{51\cdot0} \times 100 = \underline{\underline{79}}$$

---

* Rose. *Physiol. Rev.*, 1938, **18**, 109.

† Macrae *et al.* *Bioch. J.*, 1943, **37**, 225.

‡ Osborne, Mendel and Ferry. *J. Biol. Chem.*, 1919, **37**, 223.

§ Mitchell. *J. Biol. Chem.*, 1924, **58**, 873.

In another method, the amount of liver cytoplasm is estimated, in rats fed on a diet containing the protein to be studied, after previous deprivation of protein.*

For Reference

ALBANESE (1947). The Amino-Acid Requirements of Man. *Adv. in Protein Chemistry*, Vol. III, p. 227. (Acad. Press Inc., N.Y.)
ELMAN (1944). Maintenance of Nitrogen Balance by the Intravenous Administration of Plasma Proteins and Protein Hydrolysates. *Physiol. Rev.*, **24**, 372.
MITCHELL (1924). Nutritional Value of the Different Proteins. *Physiol. Rev.*, **4**, 424.
ROSE (1938). The Nutritive Significance of the Amino-acids. *Physiol. Rev.*, **18**, 109.

Fats. Besides their high calorific value and their frequent content of vitamins, there is evidence that fats are essential constituents of diet. In the form of lecithins they are indispensable components of most living cells, and although certain fats are readily synthesised in the body, it is doubtless more economical for them to be supplied in the diet. Moreover, there seems good ground for regarding small amounts of unsaturated fatty acids such as linolenic and linoleic acids, or their esters, as being quite indispensable for proper nutrition. Rats fed on entirely fat-free but otherwise complete diets developed a deficiency disease affecting the skin, and their growth and reproduction were impaired. Addition of small amounts of the acids or fats named effected cure, or prevented the disease.

7. THE INORGANIC FOODSTUFFS

The Water Balance. Some 72 per cent. of the body weight consists of water, about 5·5 per cent. in the blood, a variable amount, around 25 per cent., in extracellular fluid, and the remainder within the cells : all this water is interchangeable—as is shown by studying the dissemination of injected D_2O—in rabbits that of blood and extracellular fluid speedily (within 30 secs.), and that between blood and the whole of the remainder more slowly (within 30 mins.).† The water content of the tissues is maintained reasonably constant by equalisation of the loss and gain of water.

The water of the blood plasma, lymph and extracellular fluid may be called "free water," because its molecules can freely interchange and travel, and so can carry dissolved substances without appreciable hindrance from place to place. Its amount can be determined by injecting into the blood a known amount of some diffusible non-metabolisable substance, such as sodium thiocyanate, and, after a short interval to allow of its uniform dispersal, finding its concentration in the plasma. From such experiments the "thiocyanate space" is found to be about 17 l. for a normal adult, or about 25 per cent. of the body weight.

The remaining water is held within the cells, by reason of its being bound to colloids and confined by cell membranes ; diffusion of dissolved substances between the extracellular fluid and the intracellular water is not entirely free, but is conditioned by the properties of the plasma membrane and cytoplasm.

The osmotic pressures of the blood, tissue fluid, and of the various secretions into the alimentary canal are identical. Even when very large amounts of water are drunk there is no considerable dilution of the blood ; nor is there measurable concentration of it on deprivation, except in extreme thirst.‡ By delicate methods, however, changes can be shown to occur.§

The income of water is derived from two sources, viz. water taken in as such with food or drink, and water formed by the oxidation of foodstuffs. The output of water is by the lungs, skin, urine and fæces.

---

* CAMPBELL and KOSTERLITZ. *J. Physiol.*, 1948, **107**, 383.

† FLEXNER *et al.* *J. Biol. Chem.*, 1942, **144**, 35.

‡ HALDANE and PRIESTLEY. *J. Physiol.*, 1915, **50**, 296.

§ PRIESTLEY. *J. Physiol.*, 1915, **50**, 304. MARGARIA. *J. Physiol.*, 1930, **70**, 417.

The water formed by oxidation of foodstuffs can be calculated since we know that, from 100 g. protein we get 41 g., from 100 g. fat 108 g., and from 100 g. carbohydrate 55 g. water, and so we deduce that with average sedentary diets about 350 c.c. of water per day is formed by oxidation.

The total exchange of water per day is approximately as follows for average conditions for a sedentary adult in a temperate climate :

Income.		Output.	
Drink	1,450 c.c.	Urine	1,500 c.c.
Food	800 ,,	Skin	600 ,,
Oxidation	350 ,,	Lungs	400 ,,
		Fæces	100 ,,
	2,600 c.c.		2,600 c.c.

The proportion of the water lost by the respective channels of kidneys, skin, lungs or bowel will naturally vary with conditions. In diuresis relatively more will pass out by the urine, in sweating more by the skin, when dry air is breathed more by the lungs, in diarrhœa more by the bowel; when more passes out by one channel, there will be less to be eliminated by the other routes. The biggest variables are the volume of fluid drunk and the amount lost by sweating: when the former is increased the urine output increases to keep the balance, when the latter increases urine output is reduced and the fluid drunk increased. The specific gravity of the urine is the best objective indication as to adequacy of water intake, since the kidneys must excrete at least about 300–500 c.c. daily in order to remove the necessary substances in solution; if the salt intake is raised, however, the minimum obligatory urine volume is increased.*

In a person at rest and not visibly perspiring, the loss of water by lungs and skin can be estimated by weighing the subject at intervals; the loss of weight is mostly due to water loss, but allowance can be made for the small loss due to the difference in weight between oxygen intake and CO_2 output. Proceeding in this manner, it has been found † that the rate of loss from "insensible perspiration" is proportional to the basal metabolism, and in an average man is about 40 grammes per hour, of which 18 grammes represents water lost by the lungs, 18 grammes water lost by skin, and 4 grammes the balance between weights of O_2 intake and CO_2 loss.‡

With alterations of body weight there is usually a proportional loss of water and of solid constituents, but in some cases there is a proportionally greater loss or gain of water, which it is important to take into account. A loss greater than 20 per cent. of the body water is not usually compatible with life.

Regulation of the water intake is principally effected through the agency of thirst (Chapter XLIV.; for the conditions regulating urine secretion *v.* Chapter XLVII. and for sweating, Chapters XLVIII. and XLIX.).

Excessive water intake, if uncompensated, causes a serious condition, by lowering the osmotic pressure of the tissue fluid, which causes the cells to swell: it can prove fatal.

In abnormal states the distribution of body water is often altered. Thus, in conditions of shock, there is loss of plasma into the tissue fluid, so that "hæmoconcentration" occurs, and is shown by a rise in the hæmoglobin content of the blood. This can often be set right by intravenous injections of blood plasma: saline solutions are useless, because they do not increase the colloid osmotic pressure of the blood which is needed to enable it to retain its fluid. If for any reason a patient cannot drink, it is again wrong to give saline solution, as this only throws extra work on the kidneys in eliminating the salt, and water goes out with it: in this case 5 per cent. aqueous glucose should be given intravenously. But if there has been dehydration due to loss of secretions and tissue fluid, as in diarrhœa or vomiting, then the administration of intravenous saline would be correct.

---

* BLACK *et al.* *J. Physiol.*, 1944, **102**, 406; MCCANCE and YOUNG, *J. Physiol.*, 1944, **102**, 415.

† BENEDICT and ROOT. *Arch. Int. Med.*, 1926, **38**, 1.

‡ BENEDICT and BENEDICT. *Proc. Nat. Acad. Sci.*, 1927, **13**, 364.

Dehydration occurs much more readily in infants than in adults, because, whereas in adults the daily water exchange only amounts to about one-eighth of the bulk of extracellular fluid, in infants it is as high as one half.

For Reference

Adolph (1933). The Metabolism and Distribution of Water in Body and Tissues. *Physiol. Rev.*, **13**, 336.
Marriott (1923). Anhydræmia. *Physiol. Rev.*, **3**, 275.
Peters (1935). Body Water. (Baillière, Tindall & Cox.)
Peters (1944). Water Exchange. *Physiol. Rev.*, **24**, 491.
Rowntree (1922). The Water Balance of the Body. *Physiol. Rev.*, **2**, 116.
Van Slyke (1926). Factors affecting the Distribution of Electrolytes, Water and Gases in the Animal Body. (Philadelphia and London : Lippincott & Co.)

Saline Constituents. Mineral salts are essential to life. If an animal be fed with a diet from which all the salts have been removed, it loses appetite, becomes ill, and dies sooner than if completely starved. An animal in a state of salt-hunger becomes very weak, and shows signs of paralysis ; food, still undigested after prolonged periods in the stomach, is vomited up, and it is clear that the functions of the entire body are profoundly upset.

In salt-hunger the body is continually giving off inorganic constituents in the urine. The amount of these is smallest when the animal is supplied richly with organic foodstuffs. It seems that certain inorganic ions are held in a state of combination with the tissue constituents.* If the food supplied is insufficient, the animal lives on its own tissues, thus setting free salts which appear in the urine. An interesting condition is miner's cramp, met with in workers who drink water and perspire freely. So much sodium chloride is lost in the sweat that the urine, though small in bulk, is sometimes quite free from chloride. (Normal urinary Cl = 7 — 10 g./l. as NaCl.) † The cramp is due to excess of water in the tissues, and can be avoided by the subject drinking water containing 10 grains of salt per gallon (0·105 per cent.).

A large number of elements are probably needed for the maintenance of health and normal growth, and these are required by the body in very different amounts. The daily adult intake for some of the more important may be put approximately at :

Calcium	0·7 g.
Sodium	6·0 ,,
Potassium	4·0 ,,
Iron	12 mg.
Phosphorus	1·25 g.
Iodine	0·05 mg.
Chlorine	20·0 g.
Zinc	12 mg.

Most of these elements are taken in excess of the minimal requirements, any excess over this being excreted, and it is often uncertain what the minimal requirement actually is. Usually a state of balance is maintained, as in the following example for Ca, Mg and Fe :

Balance Sheet for Calcium, Magnesium and Iron.

(Amounts for six subjects over fourteen days (McCance and Widdowson.‡))

	Intake from Food, g.	*Output, g.*			*Balance, g.*
		Urine.	*Fæces.*	*Total.*	
Calcium	61·68	13·69	48·99	62·68	— 1·00
Magnesium	22·80	9·28	13·30	22·58	+ 0·22
Iron	0·6033	0·0078	0·607	0·6148	— 0·0115

---

* Baird and Haldane. *J. Physiol.*, 1922, **56**, 259.

† Hancock, Whitehouse and Haldane. *Proc. Roy. Soc.*, 1929, **105B**, 43.

‡ McCance and Widdowson. *J. Physiol.*, 1938, **94**, 148 ; *Bioch. J.*, 1939, **33**, 523.

Sodium, chiefly in the form of chloride, plays an important part in the maintenance of the proper osmotic pressure of the body fluids, and the sodium ion is also essential in proper concentrations in those fluids. Sodium deficiency is rare in Europe, but is seen after removal or disease of the suprarenal cortex.

Potassium is an essential ion in all cells, and deficiency of it is very rare.

Iron is essential for the formation of hæmoglobin and of the various hæmatin-like compounds, such as cytochrome, which play a very important part in carrying on the oxidative processes in the body. Much of the iron taken by the mouth is unabsorbed. Milk is poor in iron (about 1 mg./l.), but the infant is born with a store of iron in the liver sufficient to last it for several months. Good sources of iron are meat, potatoes, liver, kidney, watercress and most vegetables. Iron compounds are probably reduced to the ferrous state in the alimentary canal, and unabsorbed iron is excreted as ferrous sulphide. Deficiency of iron leads to anæmia, but the taking of excess of iron does not lead to formation of larger amounts of hæmoglobin than normal.

Very little iron is excreted per day, the iron store of the body being kept up by conservation of iron. When the hæmoglobin of the red cells is disintegrated on their destruction, the iron is stored for a time as a ferric hydroxide complex with protein, known as *hæmosiderin*, in the liver, and is re-utilised as demand arises for the construction of fresh hæmoglobin.

Copper in a very small amount appears to be necessary for the assimilation of iron and formation of hæmoglobin.

Cobalt is a constituent of the hæmopoietic principle of the liver.

Manganese is another element which must be supplied in minute amounts. In its absence the processes of reproduction are disturbed.

Magnesium is also essential. It is chiefly present in the body in the bones, as a phosphate. A rat requires not less than 5 mg. of Mg per 100 g. food. The human requirement is about 0·6 g./day. In magnesium deprivation there are vascular disturbances, hyperexcitability, convulsions and death.* The effects of Mg. deficiency are intensified if the diet is very rich in Ca.† The Mg. ion acts antagonistically to the Ca ion with regard to the excitability of tissues.

Chloride is the chief anion of the body, but especially of the interstitial fluid: it is readily absorbed as soluble chlorides. In shortage of chloride the ion is tenaciously retained in the body, and the urine, which normally carries the excess of ingested chloride over the requirements, may then become totally chloride-free.

Iodine is a constituent of thyroxin, the hormone of the thyroid. About half of the body iodine is present in the thyroid. Iodides are rapidly absorbed and are excreted unchanged.

Bromide is also quickly absorbed, and is normally present in traces in the body, where it mingles with the chlorides.

Calcium ‡ is most abundant in milk and cheese; eggs and green vegetables contain some, but meats, fish and bread are poor in calcium. Some of the calcium requirement can be obtained by drinking hard water. Much of the calcium of the diet is not absorbed, the fraction so remaining depending on the other constituents of the diet, and on the extent to which the duodenal contents retain an acid reaction. Bile salts and sugars aid in the absorption of calcium, while alkalies or large amounts of phosphates retard it by formation of the insoluble calcium phosphate. Fats also retard it. Growing children and pregnant women and nursing mothers need more calcium than other persons, and about 1·6 g./day is suggested for them.

Calcium is essential for the formation of the bones and teeth, for blood clotting, and for the maintenance of the normal excitability and permeability of different types of cells.

---

* Kruse, Orent and McCollum. *J. Biol. Chem.*, 1932, **96**, 519.

† Tufts and Greenberg. *J. Biol. Chem.*, 1938, **122**, 715.

‡ Leitch. *Nutr. Abstr. Rev.*, **6**, 553.

The calcium of the blood is practically all in the plasma (or serum); its amount varies with the amount of plasma protein and usually is about 10 mg./100 c.c. Not all of it is diffusible or ionisable but usually only about half of the total amount. Practically all that is dialysable is fully ionised. The fraction that is non-dialysable is attached to protein and in fact the calcium ion concentration of the plasma can be calculated from the formula of McLean and Hastings :—

$$\frac{[Ca^{\cdot\cdot}] \times [\text{Protein}]}{[\text{Ca Proteinate}]} = 10^{-2.22} \text{ at } 25° \text{ and } pH\ 7{\cdot}35.$$

A nomogram for finding the serum calcium from the total Ca and total protein is given in Fig. 530.

The ionic calcium also varies in the opposite direction to the *p*H of the plasma, but no strict relationship between them seems to have been yet discovered. The Ca of the cerebro-spinal fluid is all in the ionic state, and as a rule equals that of the ionic calcium of the blood plasma.

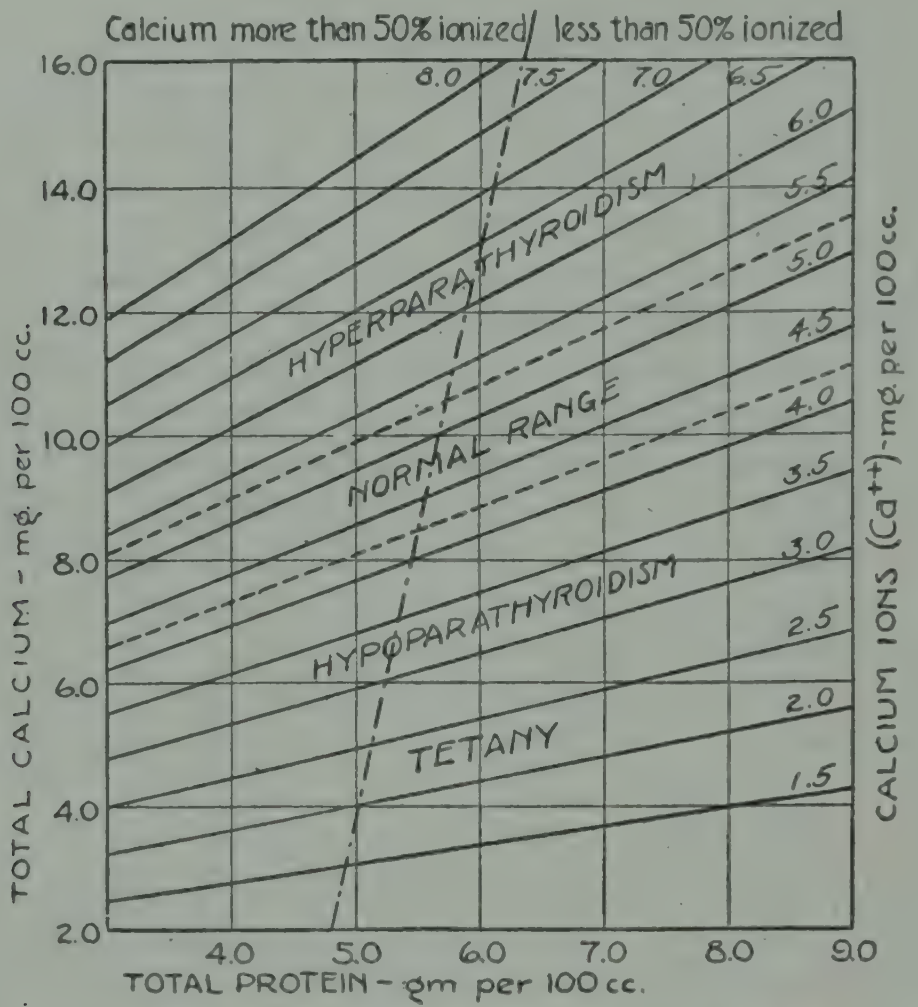


FIG. 530. Nomogram for finding $Ca^{\cdot\cdot}$ of serum from total Ca and protein. The position of the point of intersection of any pair of rectangular co-ordinates with respect to the more horizontal oblique lines gives the $Ca^{\cdot\cdot}$ of the serum, *e.g.* at 8 per cent. protein and 12 mg. Ca/100 c.c., the $Ca^{\cdot\cdot}$ is 5 mg./100 c.c. (McLEAN and HASTINGS, *Amer. J. Med. Sci.*, 1935, **189**, 601.)

Soluble calcium salts, such as chloride, lactate or gluconate, given by mouth, cause an increase in the calcium of the plasma in an hour or two ; a few hours later this returns to normal again. Amounts varying from 10 to 50 per cent. of the intake of calcium by the mouth are excreted in the urine, and from 50 to 90 per cent. are excreted in the fæces and probably represent unabsorbed calcium salts. But some state that when calcium salts are given by mouth there is no increase in the calcium of the fæces ; on the other hand, it is claimed that in fasting the calcium content of the fæces is but little reduced.

Phosphorus is obtained partly in the form of phosphates of inorganic cations, but largely as organic phosphorus compounds. In the course of their metabolism in the body, these liberate phosphoric acid. Thus various phosphatides, nucleic acid derivatives, etc., all ultimately yield phosphoric acid. Part of the phosphorus present in cereals is in the form of phytin (inositol hexaphosphate), and this source of phosphorus is not utilisable in the body. Sources of available phosphorus are similar to those of calcium, but meat and fish, though poor in calcium, contain plenty of phosphorus.

Phosphorus, in the form of various derivatives of phosphoric acid, plays an important part in the processes of intermediary metabolism. Thus sugars and fats are metabolised through the intermediation of phosphorylated derivatives.

Phosphorus and calcium should be provided in a proper ratio, which in most human diets is about 2 P to 1 Ca. In bone the ratio is 2 Ca : 1 P, and that ratio is favourable for growing animals. Since both are principal constituents of bone and teeth, they must often be considered together.

8. BONE AND BONE FORMATION

Bone is a living tissue which consists of 25 per cent. water and 30 per cent. of an organic matrix impregnated with inorganic material. When a bone is soaked in dilute acid the inorganic material is dissolved away, and the remaining organic matrix is elastic and flexible. Calcination, by removing the organic matrix, leaves only the inorganic material as a friable residue of bone ash, forming some 45 per cent. of the weight of the original bone. The chief organic constituent is a collagenous protein called ossein. The inorganic material consists chiefly of calcium phosphate and carbonate, with some magnesium phosphate. Examination of the X-ray spectrum of bone shows it to be crystalline in structure, and to correspond to one of the minerals known as apatite or hydroxyapatite, *i.e.* either

$$3\ Ca_3(PO_4)_2.CaCO_3 \text{ or } Ca_3(PO_4)_2.CaOH_2.$$

It is not proposed to describe here the histological features of bone formation and in Chap. XLII. we shall deal with some features of bone formation in relation to vitamins. That the cartilaginous basis and the growing bone are formed by the osteogenic cells of the periosteum and endosteum has been shown by the direct method of tissue culture. Cartilage cells, or bone cells, when once formed, however, multiply as such, though the osteogenic layer of the periosteum can produce either cartilage cells or bone cells.

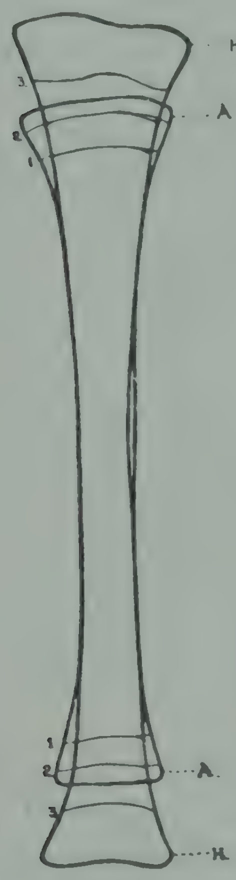
The foundation of cartilage on which a long bone is subsequently laid serves to determine roughly what the shape of the finished bone will be. It is, in fact, possible to cultivate *in vitro* a portion of undifferentiated mesenchyme from the thigh region, and to witness the emergence of a roughly formed femur in tissue quite removed from its usual surroundings. The final detail of the architecture of the bones and of the joints is, however, the result of the operation of various extrinsic factors. Among these, some are mechanical, *e.g.* due to local stresses and strains, and others are chemical. Thus, the growth and quality of the skeleton is largely influenced by such endocrine organs as the anterior pituitary, thyroid, parathyroid and sex organs, the hormones of the latter in particular being responsible for the difference in form and growth of the skeleton in the two sexes.

In the development of the long bones, the primary spongy bone which at first replaces the scaffolding of cartilage is itself only provisional, and is reabsorbed by the activity of " osteoclasts," the while new bone is being laid down to replace it. This simultaneous absorption and formation of bone is not confined to fœtal life or even to childhood, but continues, though at a slower tempo, throughout life, so that bones are far from being the static and finished structures they are commonly believed to be.

As a bone grows it increases in thickness by absorption from within the marrow cavity while sub-periosteal bone is being laid down on the outside. Hence, a layer of bone which at an early stage lies just beneath the periosteum comes at later stages to appear to sink into the underlying bone, until it finally reaches the layer adjoining the medullary cavity, after which it disappears. This can be shown by feeding a growing animal for a time on madder. This colours the new bone red, and on cross section of the shaft of a long bone from the animal, the new bone deposited during the period of madder feeding stands out clearly as a coloured ring. If the animal is killed at or shortly after the period of madder feeding, the coloured ring is at the surface of the shaft, but as the interval between feeding and killing is lengthened, so the position of the coloured ring is found to lie further and further away from the surface. This method of growth explains why it is possible to put a fœtal femur inside the medullary cavity of an adult femur, although we know that one has developed by a process of continuous growth from the other.

Long bones grow in length by progressive growth and ossification of the cartilaginous epiphyseal plate. During certain illnesses the rate of growth is greatly slowed, and the lamina of bone then laid down is more dense than normal, and remains so, as a permanent record visible in an X-ray of the bone. As this lengthening occurs it is accompanied by a good deal of re-moulding of the outer surface of the epiphysis by sub-periosteal reabsorption (Fig. 531).

Since the bones and teeth stand out from other tissues by reason of their high content of calcium salts, much attention has been directed to finding how this calcification has been brought about. It seems probable that the essential feature of the process consists in the construction, first of an organic matrix by the osteogenic cells, and then the calcification of this matrix by the formation in it of a precipitate of calcium phosphate. The precipitated calcium phosphate then appears to adsorb ions of $Ca^{\bullet\bullet}$, OH', and CO_3' to form the apatite-like bone salt.*



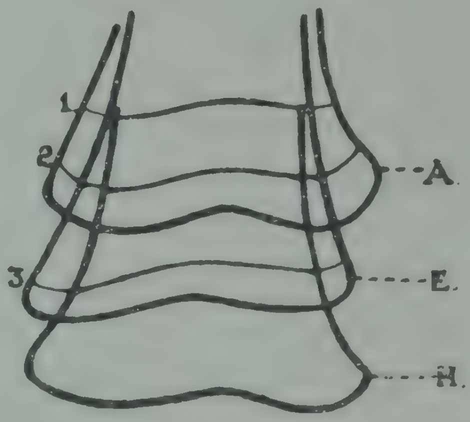


Fig. 531. Superimposed radiographs of bones of a child taken at intervals over a period of 106 weeks. (Radiographs by H. A. Harris.) On right, outlines of three pictures of lower end of the femur. Between A and E, 63 weeks elapsed, and between E and H, 43 weeks. The lines of growth retardation corresponding to periods of illness are at 1, 2 and 3. On left, A and H with 106 weeks' interval, are of tibia of same child. Note how much bone has been re-absorbed to form the final shape of H in both cases. (Brash, *Edin. Med. Journ.*, 1934, **41**, 305.)

Processes of calcification and of decalcification are apparently always going on in bone—sometimes, as in growing bone, calcification greatly exceeds decalcification, and sometimes, as when there is present an excess of the hormone of the parathyroid, decalcification is in excess of calcification. The two processes are probably independent, and occur locally as a result of cell activity, and in the normal adult a state of equilibrium between the two processes occurs. That continuous replacement of the inorganic material of bone goes on is shown by the fact that when salts of radio-active calcium, or phosphates containing radio-active P are administered, the radio-active element can be detected in all the bones in a few hours or days.

The deposition of calcium phosphate in the formation of bone is probably bound up with the fact that the formation of a precipitate of an ionisable substance from a solution depends on the concentration of the constituent ions reaching a value at which their product exceeds the "solubility product" of the substance concerned. In this case, when the solution is in equilibrium with a small amount of solid calcium phosphate, *i.e.* when the solution is saturated, then $[Ca^{\bullet\bullet}] \times [PO_4{}''']= K$. Protein, by forming unionised Ca proteinate, decreases the amount of Ca in solution and, as we have seen (p. 814), the concentration of Ca ions in plasma can be calculated if the total Ca and total protein are known. If the concentration of either the $Ca^{\bullet\bullet}$ or the $PO_4{}'''$ in the saturated solution is increased, then calcium phosphate is precipitated until the product $[Ca^{\bullet\bullet}] \times [PO_4{}''']$ falls to the solubility product value. Precipitation of calcium phosphate from a saturated solution can thus be brought about by adding a soluble phosphate (in just the same way as we can throw down sodium chloride from its saturated solution by the addition of HCl), or alternatively by the addition of excess Ca ions in the form of a soluble calcium salt.

---

* Logan and Taylor. *J. Biol. Chem.*, 1937, **119**, 293.

Now the product of the $Ca^{\bullet\bullet}$ and $PO_4{}'''$ ions in normal blood plasma is about equal to the solubility product of Ca phosphate, and therefore an increase in say, $PO_4{}'''$ ions will throw down calcium phosphate ; moreover, the calcium phosphate thrown down will not be redissolved unless the ionic product falls far below the solubility product.

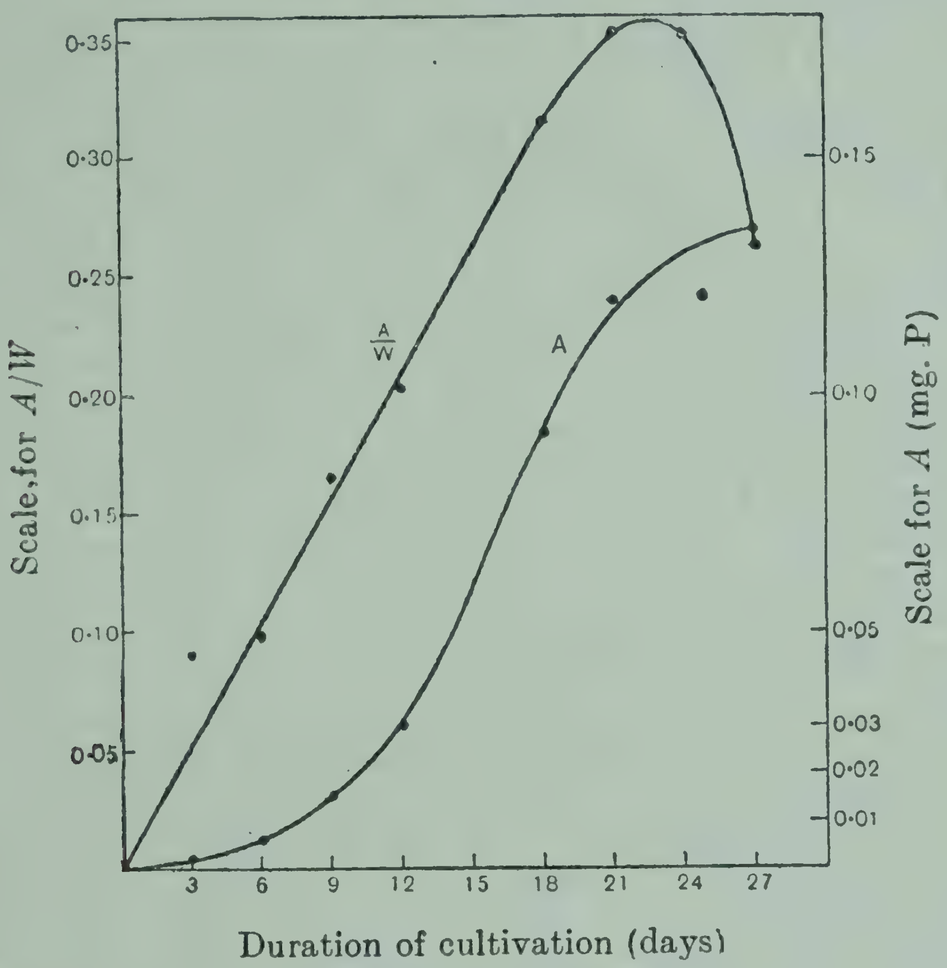


Fig. 532. Production of phosphatase during development of femora of fowl embryos excised at sixth day.

A = Phosphatase per femur (mg. P hydrolysed from glycerophosphate).
A/W = Phosphatase per mg. dry weight of femur (after Fell and Robison. *Bioch. J.*, 1929, **23**, 767).

It was shown by Robison that in the process of calcification an important part is played by an enzyme called *bone phosphatase*, which is produced by the osteoblasts. This enzyme can decompose phosphoric acid esters (such as glycerophosphate or hexose monophosphate), which are unionised, with liberation of $PO_4{}'''$ ions, and so lead to such a local increase in the ionic product as to exceed the solubility product of calcium phosphate, which is therefore deposited from the blood plasma. The enzyme is present in ossifying (but not in non-ossifying) cartilage : similar phosphatases also occur in other tissues, especially in the small intestine and kidney ; we cannot here discuss the interesting question as to why these last two tissues do not ossify, except to state that the bone matrix evidently has something to do with the deposition which occurs in bone. Bone salts are deposited on a piece of boiled ossifying cartilage placed in the abdominal cavity.

Phosphatase is not originally present in the cells of the mesenchyme from which the bone is to be formed, and the appearance and increase of phosphatase in a developing bone can be followed by making estimations of phosphatase on portions of a bone growing *in vitro* (Fig. 532).

It is perhaps significant that magnesium, which is also an invariable constituent of bone, greatly accelerates the action of phosphatase.

JOINTS

A joint is a tissue space lined by modified connective tissue cells. The articular surfaces are covered by hyaline cartilage and are kept in contact mainly by the supporting structures of the joint, such as the ligaments. The joint capsule is entirely of mesenchyme origin, and the inner layer of the synovial membrane is provided with blood vessels, lymphatics and nerves, and lined by large modified connective tissue cells which simulate an epithelium. The intermittent pressure to which a joint is subjected is important in keeping the articular surfaces in normal condition, and no doubt also in facilitating exchanges between the blood and the synovial fluid. The synovial membrane, unlike the articular cartilage, has high powers of regeneration if damaged, and is difficult to stretch.

THE SYNOVIAL FLUID is normally small in amount, *e.g.*, about 0·5 c.c. in a normal human knee-joint. It is a clear, pale yellow fluid, viscous and slippery to the feel, owing to the presence of mucin, and is subjected when *in situ* to slight if any positive pressure when the joint is at rest. It contains a number of cells, in man about 60 per c.mm., though the number and type of cells varies considerably, no doubt as a result of the occurrence of slight degrees of trauma to the joint. Of the cells present in normal fluid about half are monocytes and a quarter lymphocytes ; the rest are polymorphs, and other cells similar to those of blood.

In composition, the fluid is distinguished by the presence of mucin, about 0·14 per cent., and it contains about 2 per cent. of other proteins. Inorganic ions are nearly the same as in the plasma, such difference as exists being what would be expected from the Donnan equilibrium if we regard the synovial fluid as being in equilibrium with the blood. The high Ca content is at variance with this, however. Other constituents, such as urea, uric acid, etc., are at the same concentration as in the plasma. The relatively low protein content, as compared with the plasma content of 8 to 9 per cent., together with the fact that the albumin/globulin ratio is 3·9 as compared with 1·1 in the serum, means that, as in other body fluids, the passage of proteins from the blood stream is relatively difficult. It contains no fibrinogen and so does not clot.

There have been many suggestions as to how the synovial fluid is formed, but little experimental work. The most likely view is that the fluid is the liquid matrix of connective tissue, that the mucin originates therein, and that the fluid in this tissue space is in diffusion equilibrium with the blood. It has been abundantly demonstrated that substances of small molecular size pass readily from blood to synovial fluid, or *vice versâ*, and it is easy to show that acid dyes of high diffusibility pass freely from the blood into the joint.* Particulate matter, such as carbon particles, introduced into a joint, are slowly phagocytosed and removed by the lymphatics.

For Reference

BAUER, ROPES and WAINE (1940). The Physiology of Articular Structures. *Physiol. Rev.*, **20,** 272.

ELVEHJEM (1935). The Biological Significance of Copper and its Relation to Iron Metabolism. *Physiol. Rev.*, **15,** 471.

FENN (1940). The Rôle of Potassium in Physiological Processes. *Physiol. Rev.*, **20,** 377.

GARDNER and PFEIFFER (1943). Influence of Estrogens and Androgens on the Skeletal System. *Physiol. Rev.*, **23,** 139.

VON GETTINGEN (1935). Manganese, etc. *Physiol. Rev.*, **15,** 175.

HUGGINS (1937). The Composition of Bone and the Function of the Bone Cell. *Physiol. Rev.*, **17,** 119.

LOGAN (1940). Recent Advances in the Chemistry of Calcification. *Physiol. Rev.*, **20,** 522.

SCHMIDT and GREENBERG (1935). Occurrence, Transport, and Regulation of Calcium Magnesium and Phosphorus in the Animal Organism. *Physiol. Rev.*, **15,** 297.

SHOHL (1923). Mineral Metabolism in Relation to Acid-base Equilibrium. *Physiol. Rev.*, **3,** 509.

SHOHL (1940). Mineral Metabolism. (Chapman and Hall.)

STEWART and PERCIVAL (1928). Calcium Metabolism. *Physiol. Rev.*, **8,** 283.

---

* ENGEL. *Q. J. Exp. Physiol.*, 1940, **30**, 231.

CHAPTER XLII

THE VITAMINS

If we feed an animal on a mixed diet containing proteins, fats, carbohydrates and salts in the proper proportions, but in a state of purity, *e.g.* a mixture of caseinogen, starch, sugar, salts and lard, we find that the animal becomes ill, ceases to grow if young, and finally dies. This had already been pointed out in 1880 by Lunin, but the matter was taken up again by Hopkins,* who worked on young growing animals. He found that in the case of rats it was sufficient to add 3 c.c. of milk daily to the artificially compounded diet for growth to take place normally (Fig. 533). Three cubic centimetres of milk contain only 0·08 g. total solids, of which the greater part consists of substances already present in the artificial diet; yet this small amount of fresh milk enabled the rats to add 0·5 g. a day to their weight. These observations were the beginning of numerous series of experiments by many different observers. It has been found that practically all fresh foods contain extremely small traces of essential organic substances to which the name "vitamins" has been given. Each of them is a catalyst indispensable for the occurrence of some metabolic reaction, but is not found in the body and must therefore be provided in the diet. They are essential for the utilisation of the food, for the maintenance of health and for growth. Fifteen or more vitamins have been clearly distinguished according to their different properties and mode of action. It should be stressed, however, that "except in infancy and childhood, there are but few instances where good health cannot be maintained by eating properly

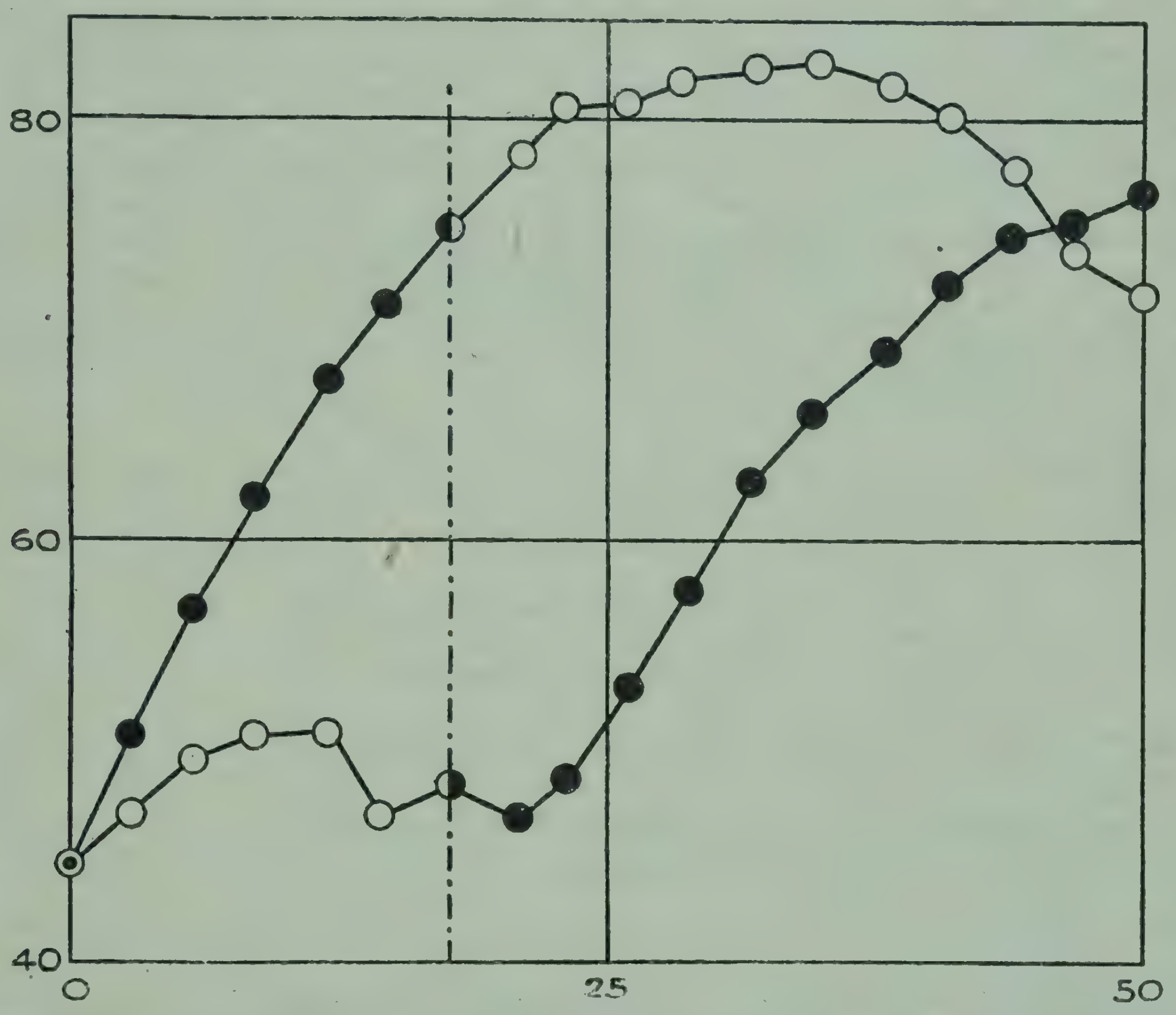


Fig. 533. Two Curves showing Growth of Rats with and without "vitamin A."

Lower curve (up to 18th day), eight male rats on pure dietary; upper curve, eight similar rats taking 3 c.c. of milk each a day. On the 18th day, marked by vertical dotted line, the milk was transferred from one set to the other. Average weight in grammes vertical; time in days horizontal. (F. G. Hopkins.)

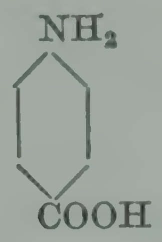
---

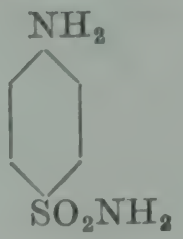
* Hopkins. *J. Physiol.*, 1912, **44**, 425.

chosen food without the constant administration of vitamin preparations" (E. Mellanby).

Several vitamins have been proved to play a definite rôle as catalysts in connection with enzyme actions concerned in intermediary metabolism.

In carrying out this rôle they enter into transitory combination with some part of the reacting enzymes, and an important consequence of this arises in the phenomenon of *competition* by analogous substances. As an example of this we may take a classical instance concerned with the processes of intermediary metabolism of certain bacteria, for which the substance *p*-aminobenzoic acid is an indispensable substance. Now it would be thought that the rather similar substance sulphanilamide might do instead, but far from this being the case its presence has the same effect as *absence* of the essential *p*-aminobenzoic acid : this is because it enters into combination with the enzyme with which *p*-aminobenzoic acid should combine, and so excludes the latter from its essential functions in intermediary metabolism, by competition.


p-*aminobenzoic acid*


Sulphanilamide

The amounts of the various vitamins needed can be only approximately stated, because they vary with the total Calorie needs. Hence, for example, vitamins A, B and C are needed in greater amount in hyperthyroidism, or when heavy work is done. In experiments on the vitamins, animals are fed for several weeks on a diet deficient only in the vitamin concerned, and are compared with controls fed with the diet, but containing the vitamin. It is important to remember, however, that certain vitamins can be produced in the intestine by the activity of bacteria.*

The vitamins are often distinguished by letters. Some are soluble in fats and occur naturally together in association with them, while others are water-soluble.

Some Important Vitamins

Fat-Soluble Vitamins

A $\begin{cases} A_1 \text{ Also termed Axerophthol, the anti-xerophthalmic or anti-infective factor.} \\ A_2 \text{ Closely related to } A_1\text{, but of small importance.} \end{cases}$
D_2 (Calciferol) or irradiated ergosterol.
D_3 Natural vitamin D.
E (Tocopherol). Anti-sterility factor.
K Essential for formation of prothrombin of blood.

Water-Soluble Vitamins

B_1 (Thiamin ; aneurin). Anti-neuritic. Beri-beri preventive : forms co-enzyme of pyruvate carboxylase.
B_2 complex, comprising :—
    Riboflavin. Component of oxidising systems.
    Nicotinamide (Niacin). Anti-pellagra factor.
    Folic acid (vitamin M). Hæmopoietic factor.
    † Biotin (vitamin H).
    † Choline.
    ‡ Pyridoxin (vitamin B_6). Rat anti-dermatitis factor.
    ‡ Pantothenic acid.
    ‡ Inositol.
    ‡ *p*-aminobenzoic acid.
B_{12}, the hæmopoietic factor of liver.

---

* Elvehjem. *Feder. Proc.*, 1948, **7**, 410.
† = significance for man doubtful.
‡ = action only demonstrated on lower species.

C (Ascorbic acid). Anti-scorbutic factor.
P Citrin or Hesperidin. A factor associated with ascorbic acid and believed to supplement its action.

FAT-SOLUBLE VITAMINS

Vitamin A_1: its absence leads to failure of growth (Fig. 533), excessive keratinisation of the skin, including an affection of the cornea called xerophthalmia, a general lowering of resistance to infection and ultimately to death. Prolonged deficiency of this vitamin during growth also causes extensive degeneration in the central nervous system, especially in the afferent nerves and tracts * due to excessive bone growth in certain positions. Whereas vitamin D controls the calcification of bone, vitamin A co-ordinates the activity of the osteoblasts and osteoclasts, so that in its absence there is loss of normal bone shape and also of co-ordination of its growth in relation to that of the central nervous system. For instance, the growth of the basi-occipital bone is abnormal, because whereas normally there is a layer of active osteoclasts at surface 1 (Fig. 534) which absorbs bone, and a layer of active osteoblasts at surface 2 laying down bone, in vitamin A deficiency the state of affairs is reversed, while at surfaces 3 and 4 normal growth proceeds. Hence the bone becomes thickened and the marrow cavity enlarged.

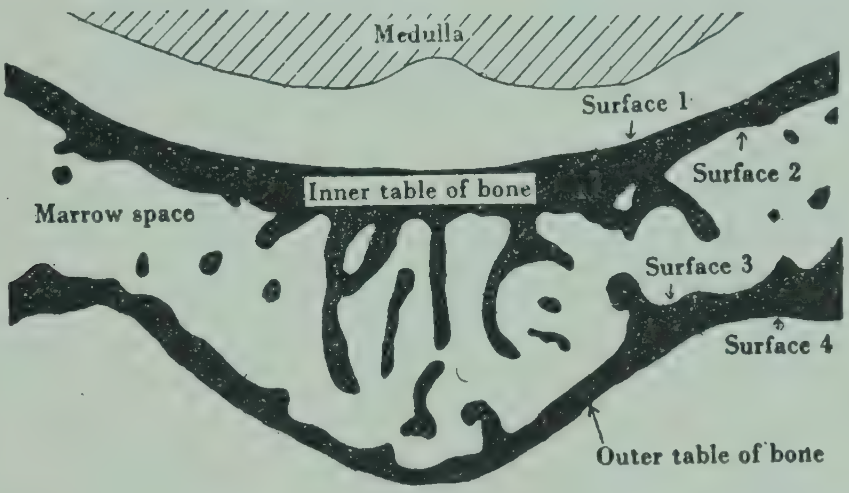


FIG. 534. Diagram of the basi-occipital bone, showing surfaces at which osteoclasis and bone formation occur. (MELLANBY. *J. Physiol.*, 1947, **105**, 382.)

On restoring vitamin A there is a return to the normal distribution of the responsible cell elements, and hence, if not too late, a reversal of the effects of the deprivation. Many other abnormalities of the nervous system may result from the same cause, including deformity of the cerebellum and medulla, increased intracranial pressure, internal hydrocephalus, and the squeezing of the whole of the central nervous system in its bony covering. Another effect of vitamin A deficiency is night-blindness, *i.e.* the inability to see in dim light. This is a consequence of reduction of visual purple in the retina owing to structural changes in the rods,† and also perhaps to a chemical relationship between visual purple and vitamin A. The occurrence of night-blindness is a convenient and early symptom of A_1 deficiency.

Vitamin A circulates in the blood, the plasma of which normally contains about 120 I.U. per 100 ml. It can be stored in the body, so that the effects of its absence from the food are not perceptible in adults till after many weeks. The normal human liver contains about 500,000 I.U., or enough to supply requirements for up to 400 days. In isolating it, the fats can be removed by saponification in the absence of oxygen.‡ The residue can be freed from cholesterol and the vitamin persists in the fraction consisting of unsaturated alcohols and hydrocarbons, but in very minute quantities.

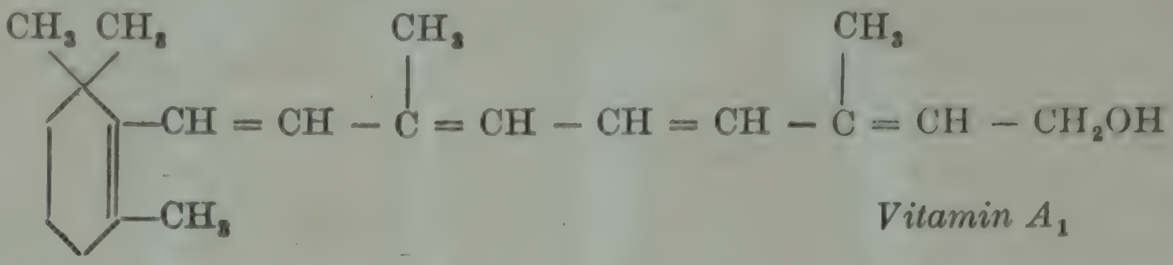
---

* E. MELLANBY. *J. Physiol.*, 1941, **99**, 467; 1947, **105**, 382.

† TANSLEY. *J. Physiol.*, 1931, **71**, 442; *Proc. R. S.*, 1933, **114B**, 79.

‡ MCCOLLUM and DAVIS. *J. Biol. Chem.*, 1914, **19**, 245; DRUMMOND and COWARD, *Lancet*, 1921 (ii), 698; DRUMMOND, CHANNON and COWARD, *Biochem. J.*, 1925, **19**, 1047.

Vitamin A_1 is a viscid colourless oil of the elementary composition $C_{20}H_{30}O$. Its structure * is :



Vitamin A_1

It is insoluble in water but soluble in all organic solvents, and related to the unsaturated hydrocarbon carotene, $C_{40}H_{56}$. This is a crystalline pigment which gives a yellowish colour to certain natural fats and oils. Both carotene and vitamin A give a blue colour when treated with antimony trichloride in chloroform, and may be estimated colorimetrically by this means. Another method is by the determination of the coefficient of absorption in the ultraviolet (325 $m\mu$).

It would seem that, although animals cannot synthesise vitamin A, they can partially convert carotene into it. The amount of the purest vitamin A which is required to keep a rat from developing symptoms of vitamin A lack is only about 0·5μg. daily, whereas about 1μg. of carotene is needed. Carotene is thus to be regarded as the pro-vitamin of A. Probably the change into vitamin A occurs in the liver, in which the greater part of it is then stored.

The international unit of vitamin A is now 0·344 μg. of a specified crystalline vitamin A acetate, dissolved in a vegetable oil to make 0·1 mg. of solution. The former standard is discarded. The standard requirement for human beings is 3,000 I.U. daily. A prolonged and extensive investigation by the M.R.C. indicates 2,500 I.U. of vitamin A or 7,500 I.U. of carotene as the daily requirement in young healthy males.† Probably many diets only supply about 1,000. Milk contains about 3–5 units per c.c. ; butter about 20–50, egg yolk about 80, spinach 100, carrots 20 units per g., and other vegetables usually much less.

Vitamin A is originally formed in the growing green plant, and it is doubtful whether any higher animals have the power of forming it *de novo* in their bodies. The cow, for instance, obtains its A vitamin from grass : after ingestion it is absorbed into the fat of the body and of the milk, so that it is taken by man largely in the form of butter or milk. The organs richest in vitamin A are liver and kidneys. It is contained in large quantities in cod liver oil, and in even larger amounts in the oils from mammalian liver. Its origin in the sea is especially in the green diatoms which serve as food for the small plankton, organisms which are the basis of the food supply of the larger fish. It is not destroyed by heat, though easily oxidised. It is absent from most specimens of lard and from many vegetable oils, such as olive oil. If, however, pigs are fed on green food, pig fat contains the vitamin. It is much more abundant in the milk of cows in summer, out at grass, than when these animals are kept in byres and fed on oil-cake, etc.

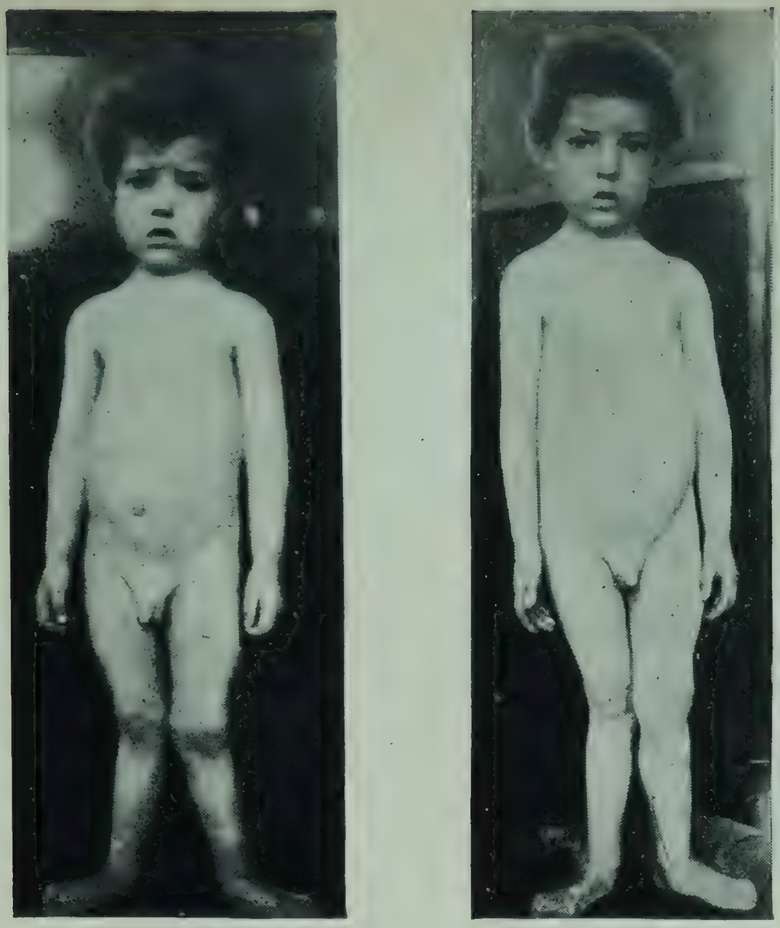
Vitamins D. The absence of growth in vitamin lack is associated with abnormal formation of bones, which gives rise to the disease of *rickets* (Figs. 535, 536), and this can be prevented by giving cod-liver oil, but not by linseed oil.‡ The substance which is present in cod-liver oil was subsequently shown to remain in the non-saponifiable fraction, and this anti-rachitic substance was called vitamin D ; unlike vitamin A it resisted the action of oxidising agents. Vitamin D also produces growth effects to a minor extent, but its chief effect

---

* HEILBRON, HESLOP, MORTON, WEBSTER, LEA and DRUMMOND. *Bioch. J.*, 1932, **26**, 1178.

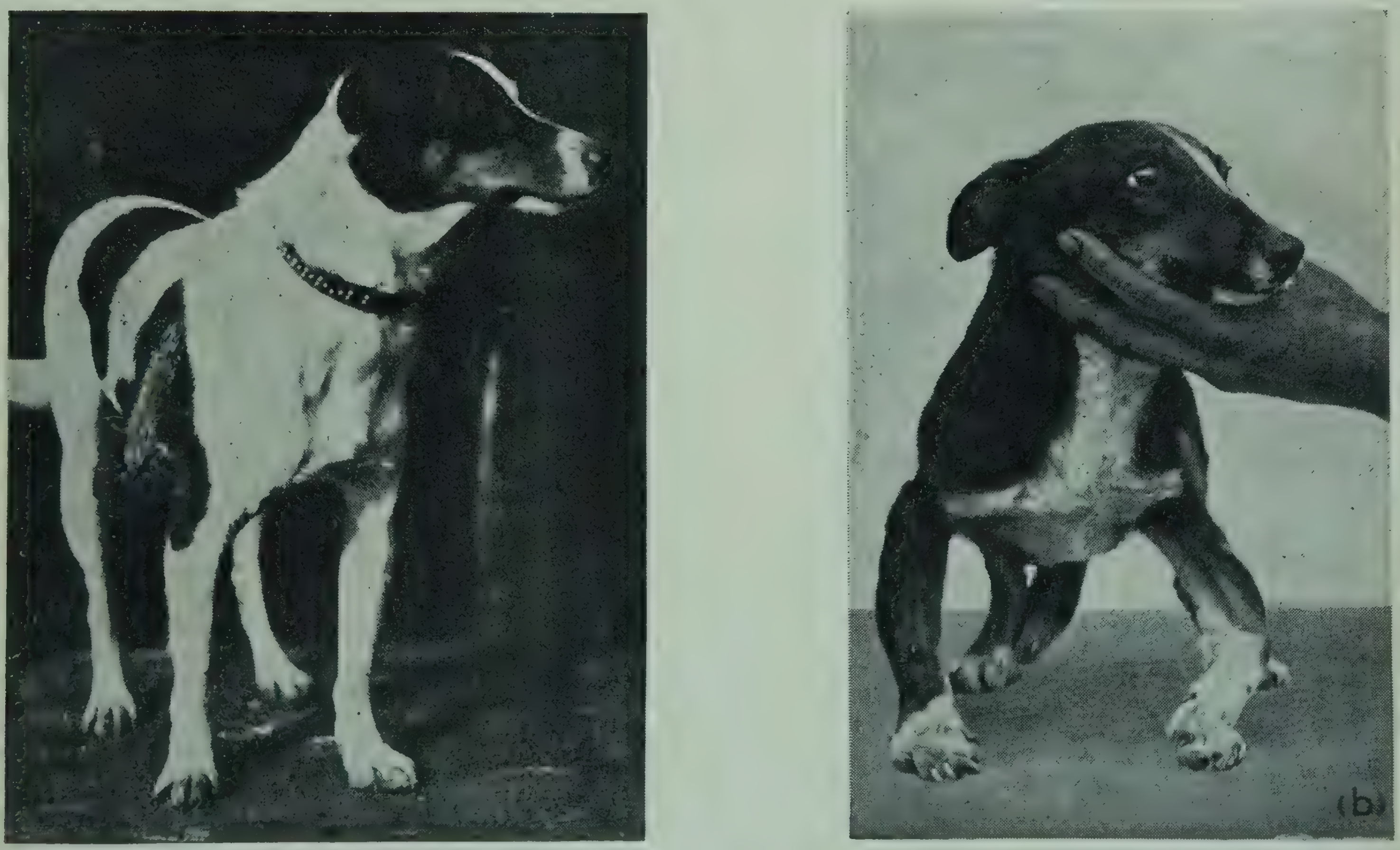
† M.R.C. Sp. Rep. No. 264 (1949). Vitamin A Requirements of Human Adults.

‡ E. MELLANBY. *J. Physiol.*, 1918, **52**, xi, liii ; *Lancet*, 1919 (i), 407.



(*a*) (*b*)

FIG. 535. Improvement in Genu Valgum by a Calcifying Diet.
(*a*) Child aged five years, with leg deformity caused by rickets.
(*b*) Improvement after six months' treatment on a highly calcifying diet.
(E. MELLANBY. "Nutrition and Disease." Oliver and Boyd, 1934.)



(*a*) (*b*)

FIG. 536.

(*a*) Normal dog reared on a good diet rich in vitamin D during a period of confinement.

(*b*) A rickety dog brought up on a diet similar to that of (*a*) except that it lacked vitamin D.

(E. MELLANBY. "Nutrition and Disease." Oliver and Boyd, 1934.)

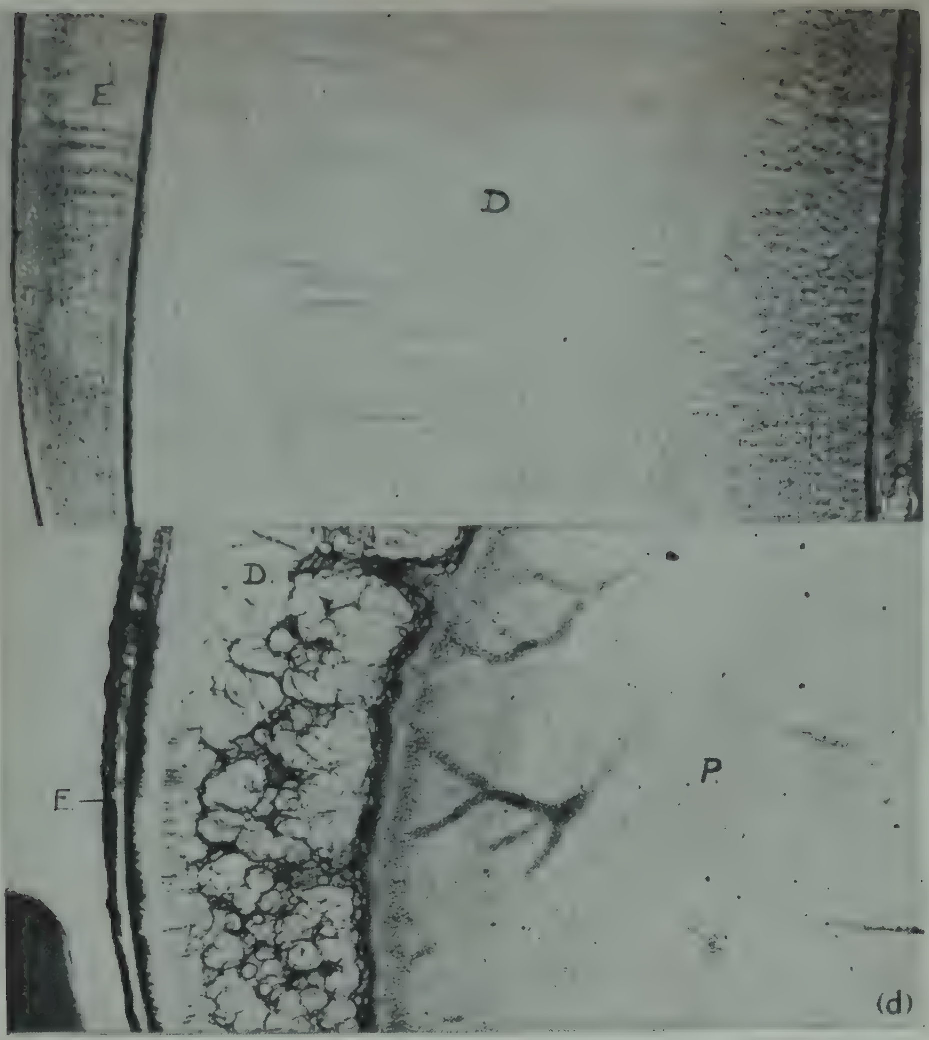


FIG. 537. Carnassials of Two Puppies.

Upper = Received cod-liver oil. Normal enamel and dentine.

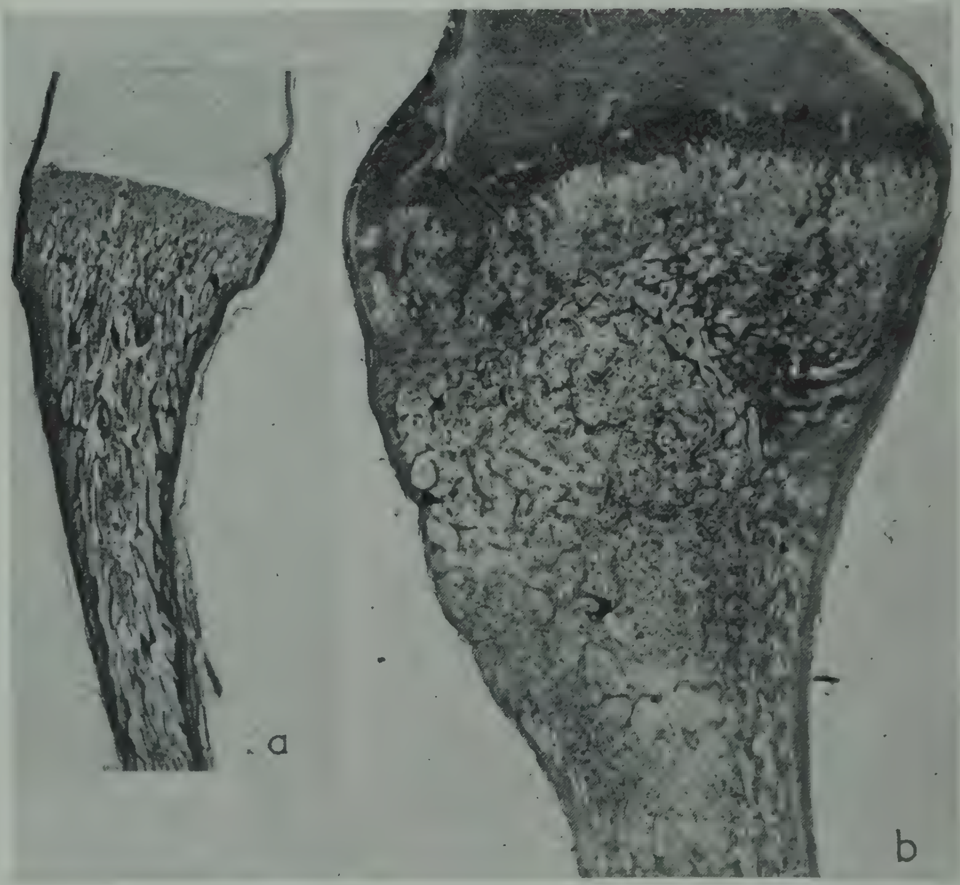
Lower = Received no fat-soluble vitamin. Thin and badly calcified enamel and dentine, and many interglobular spaces. E = enamel. D = dentine. P = pulp cavity.

(MAY MELLANBY. Taken from E. MELLANBY, "Nutrition and Disease." Oliver and Boyd, 1934.)

FIG. 538. Costo-chondral junctions of normal (*a*), and rickety (*b*) dogs.

The diets of the two dogs contained a large amount of oatmeal, which is rickets-provoking, but that of (*a*) contained cod-liver oil, while that of (*b*) contained olive oil.

(E. MELLANBY. "Nutrition and Disease." Oliver and Boyd, 1934.)



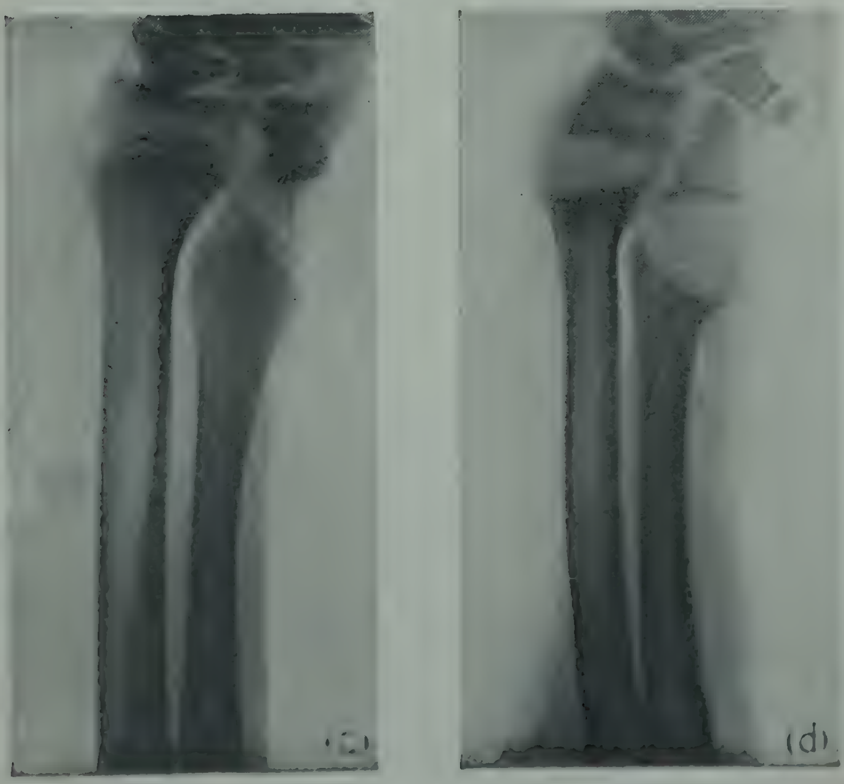


FIG. 539. Radiographs of wrists of dogs as in previous fig. (*c*) = wrist of dog (*a*) : normal. (*d*) = wrist of dog (*b*) : rickets.

(E. MELLANBY. "Nutrition and Disease." Oliver and Boyd, 1934.)

seems to be in controlling the balance of calcium and phosphorus. In normal plasma the product of the total Ca and P (inorganic) in mg./100 c.c. is over 50; in rickets both fall, the P first, and the product is below 40. The reason for this seems to be a loss of phosphate and ultimately of calcium by way of the fæces.

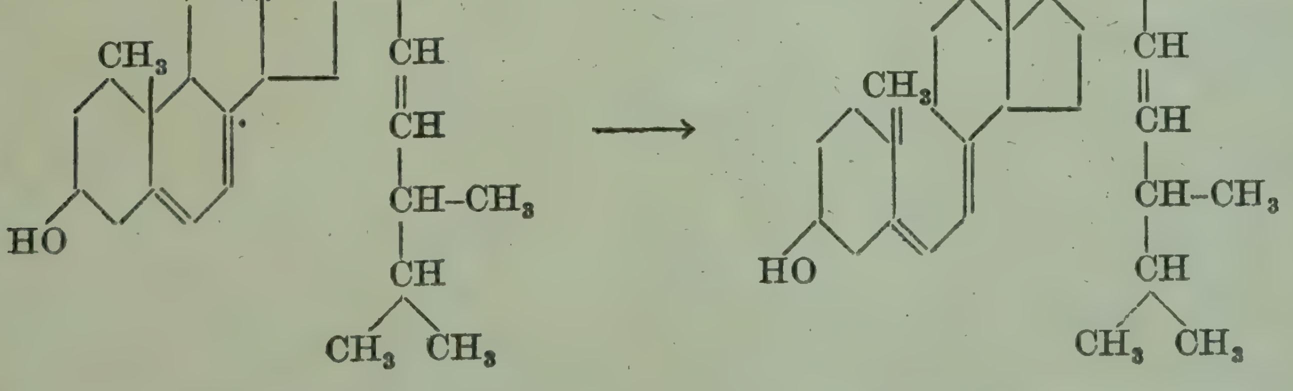
It has also been demonstrated by May Mellanby that vitamin D plays an important part in the formation of the teeth (Fig. 537). Once the dentine has been formed, however, administration of vitamin D cannot alter it either in structure or composition.[*]

The constantly growing incisor tooth of the rat is a delicate indicator of the conditions for tooth formation, the effect of replacement of vitamin D sometimes being detectable in twenty-four hours.[†]

The most marked change in rickets is a defect in bone formation normally occurring at the epiphyseal line (Figs. 538–539); the new bone at the epiphysis is inadequately calcified, soft and increased in breadth. Hence the bones show swellings at the epiphyseal junctions, and are easily deformed, resulting in bow-legs, knock-knees, etc. Ossification in cartilage consists of the following stages: (1) multiplication of cartilage cells; (2) calcification of the cartilage; (3) absorption of the calcified cartilage; (4) deposition of true bone, around cells —osteoblasts—which have invaded the calcified cartilage from the marrow. In rickets there is an accelerated multiplication of cartilage cells and a retardation in the deposition of true bone, from a failure in the process of deposition of calcium salts. If an anti-rachitic vitamin is added to the diet of a previously deprived and rachitic rat, a line of calcification soon appears transversely right across the cartilage near the epiphyseal line. This "line test" is a delicate detector of anti-rachitic property.[‡] It is evident that in the causation of rickets, or for the growth of normal bone, many factors are involved. In the first place, the presence of vitamin D is necessary. This may, however, be replaced to a certain extent by exposure to the sun's rays. Then it is necessary that calcium and phosphorus shall be present in proper amounts. An unsuitable ratio of Ca : P in the diet may also hinder the formation of bone. Thus Ca : P = 4 or 0·2 would be unfavourable; for some growing animals 2 is a good ratio. Cereals have a rickets-promoting effect owing to the fact that, calcium phytate being insoluble, and its phosphorus being present in a non-assimilable form as phytin, absorption of both $Ca^{\cdot\cdot}$ and PO_4''' are diminished.[§] It has been shown by Robison that, in the carriage of calcium salts to the bone, a part may be played by organic phosphoric esters which form soluble salts with calcium, and which have been detected in blood. Under the action of alkaline phosphatase which is present in ossifying cartilage, and which is *also present in rachitic bone*, this calcium salt is hydrolysed, and calcium phosphate is precipitated in the matrix round the osteoblasts.

The manner of action of vitamin D is unsettled, but it is probable that it primarily attacks the lipides of the tissues, with liberation of phosphoric acid,[||] the lack of which in the blood is a notable feature of rickets. It prevents loss of calcium by the bowel. As a net result, vitamin D acts by allowing an increased net absorption of calcium and phosphate, so raising their levels in the blood.[¶]

Exposure of animals to sunlight, or to ultra-violet light, can replace to a certain extent the necessity for the presence of vitamin D in the food. Investiga-



Ergosterol *Calciferol* (*vitamin* D_2).

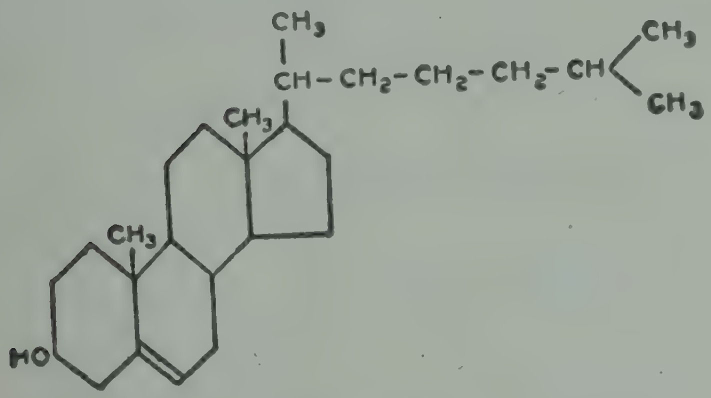
---

[*] Fish. *J. Physiol.*, 1935, **84**, 272.
[†] Irving. *J. Physiol.*, 1944, **103**, 9.
[‡] Shipley *et al.* *J. Biol. Chem.*, 1921, **45**, 343.
[§] Harrison and Mellanby. *Bioch. J.*, 1939, **33**, 1660.
[||] McGowan *et al.* *Bioch. J.*, 1931, **25**, 1295.
[¶] Nicolaysen. *Bioch. J.*, 1937, **31**, 107, 122, 323.

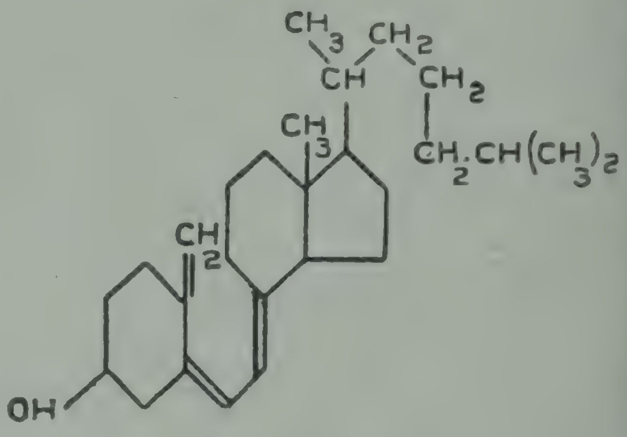
tion of the cause of the similarity between the effect of addition of vitamin D to the diet and exposure of the animal to ultra-violet light, led Steenbock to the discovery that vitamin D activity is produced by the action of ultra-violet light (250–300 mμ.) on a precursor which is present in the tissues, including the skin. It was further found* that this precursor is a sterol called *ergosterol*, $C_{27}H_{42}O$. Irradiated ergosterol has many of the properties of vitamin D.

By distillation of irradiated ergosterol in high vacuum, a crystalline isomer of double its activity was obtained; this was called "calciferol,"† or vitamin D_2. In large doses it causes loss of weight, reduced excretion of Ca and P, hypercalcæmia, toxic symptoms and death; extensive deposits of calcium salts are found in the heart, great arteries and kidneys. The anti-rachitic activity of calciferol is 40,000 international units per mg. The international unit for vitamin D activity is 0·025 μg. of a standard preperation of crystalline vitamin D_3.

It now appears that there are various natural vitamins D, the chief of which, D_3, is a very close isomer of cholesterol. It can be made by irradiation of 7-dehydrocholesterol. Its actions are similar with those of D_2.



Cholesterol.



Vitamin D_3.

The human daily requirements of vitamin D are not precisely known, but it is probable that growing children need more (*e.g.*, 2,000–5,000 units) than adults (200–500 units). Expectant mothers probably need 1,000–2,000 I.U. daily to allow of the increased Ca and P needs of fœtus and milk formation. It is essential that adequate amounts of calcium (*e.g.* 1·5g./day) and of phosphorus (2·0g./day) be also available. Poor dietaries lack these, as well as the vitamin D, and average diets are more likely to lack vitamin D than any other vitamin; hence the practice of giving it, as cod-liver oil, to children. Butter, cream and egg yolk are richest among ordinary foods; fish liver oils are especially rich (cod-liver oil averages 100 units per g.), but mammalian liver contains little.

Vitamin E. Recent experiments indicate that, for reproduction to take place, another vitamin may be necessary. Animals deprived of this vitamin may thrive and be to all appearance perfectly well, but do not produce young. Conception takes place, but the embryo fails to attain full development, dies *in utero* and becomes absorbed. The germ cells in males degenerate. There is also muscular degeneration and damage to the epithelium of the convoluted tubules of the kidney. Vitamin E is present in wheat-oil, butter and certain other fats. It is present in the unsaponifiable fraction, is not destroyed by light, heat, acids, alkalis, oxidation or hydrogenation, or distillation *in vacuo*, and has an absorption band at 294 mμ.

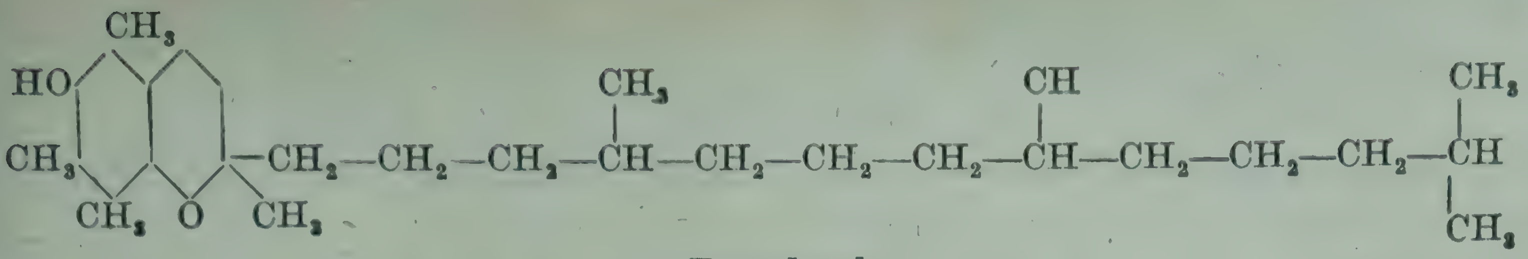
The international standard is a particular sample of racemic α-tocopheryl acetate of which 1 mg. = 1 unit. Standardisation is by determination of fertility-rate, *i.e.* the per cent. of positively mated female rats which produced a litter: the only source of vitamin E being that given in the sample, of known weight, under test.

The human requirement is probably 3 mg. per day. It possibly acts through the pituitary. There are several related components which are alcohols of the formula, $C_{29}H_{48}O_2$, called the tocopherols. The formula of α-tocopherol is :—

---

* ROSENHEIM and WEBSTER. *Biochem. J.*, 1926, **20**, 537; 1927, **21**, 389.

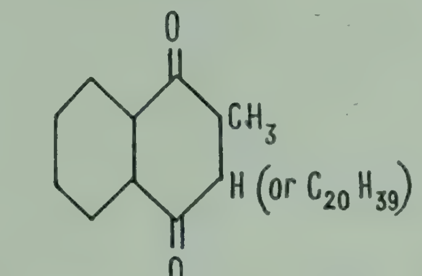
† ANGUS *et al.* *Proc. Roy. Soc.*, 1931, **108B**, 340 (and simultaneously by four other groups of investigators).



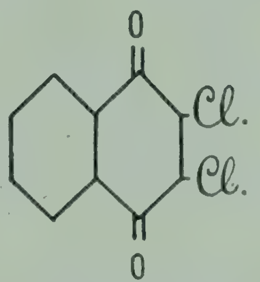
α-*Tocopherol.*

Vitamin K [*] is another fat-soluble vitamin, lack of which, in animals, produces symptoms rather like scurvy, and involving multiple hæmorrhages in skin and muscles, associated with greatly increased clotting time of the blood. It is present in many vegetable substances, especially cabbage and spinach. It is necessary for the production of prothrombin by the liver ; when it is lacking there is a lowered prothrombin content of the plasma. Bile is essential for its absorption and this accounts for the lowered coagulability of the blood in biliary obstruction. New-born infants are often deficient in vitamin K, and show increased clotting times. Chemically it is a methyl-phytyl-naphthoquinone.[†] The synthetic substitute is 2-methyl-1 : 4 naphthoquinone.

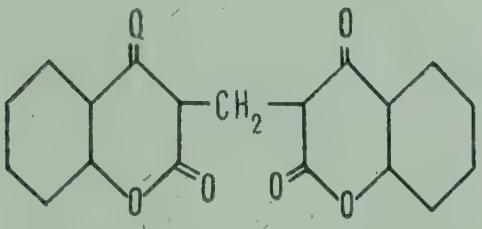
The chemically related substances dichloronaphthoquinone and dicoumarin :— act as competitors and produce symptoms like those of lack of vitamin K, *i.e.* incoagulability of the blood due to lowered prothrombin content of the plasma.



Synthetic (and Natural) Vitamin K



2.3. dichloronaphthoquinone



Dicoumarin

WATER-SOLUBLE VITAMINS

The Vitamins B. In man prolonged feeding with polished rice, *i.e.* rice from which the husks have been removed, leads to the production of the disease known as beri-beri. This is distinguished by the occurrence of polyneuritis, pains and weakness in the limbs, loss of sensation in the skin, œdema, and weakness of the heart. A somewhat similar disease may be produced by the same means in fowl or pigeons. The disease may be cured by adding the polishings to the polished rice before eating. It is evident that some substance essential to the normal nutrition of the body is removed during polishing. This was at first called vitamin B, but the effects are now known to be due to several substances.[‡] These are not stored to any extent in the body, so that animals rapidly show the effects of their withdrawal.

The individual vitamins in the complex are ten in number ; the most clearly defined has been called B_1 ; it can be differentiated from the remainder, called collectively the " B_2 complex," by their physiological effects and by the action of heat and alkali, either or both of which destroy B_1. Though both are necessary for growth, vitamin B_1, or "aneurin," or "thiamin," is believed to be related to beri-beri, while the vitamins of the B_2 complex have various functions, some of them still obscure. Yeast is rich in all, wheat germ in B_1, and milk, meat and green vegetables in B_2 complex.

---

[*] Dam *et al.* *Bioch. J.*, 1936, **30,** 1075 ; 1938, **32,** 485.

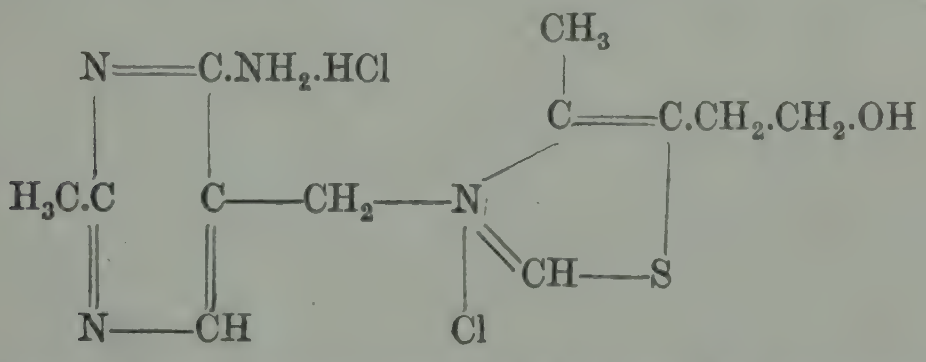
[†] Binkley *et al.* *J. Biol. Chem.*, 1939, **130,** 219.

[‡] Nelson. *J. A. Med. Assoc.*, 1938, **110,** 645.

Included in the vitamin B_2 complex are certain substances such as choline. The various factors in the complex play important parts in the control of metabolism, but these rôles are, as yet, not completely clarified. Other vitamin B_2 components in an average day's diet are said to be pantothenic acid (10 mg.), pyridoxine, biotin (0·25 mg.), inositol (1 g.) and folic acid (1·5 mg.). Whether the first three are needed by man is not known. The various members of the B_2 complex are essential for the growth of certain yeasts and bacteria, and may, in fact, be estimated by utilising this property.

Vitamin B_1 (Thiamin or Aneurin). Absence of B_1 vitamin leads in rats to wasting, failure of growth, atrophy of the testes, and sterility. The wasting is due to lack of appetite and to digestive disturbances. Secretion of gastric juice is reduced.* Among the effects of vitamin B_1 is a quickening of the rate of the heart-beat, which is slow in its absence.† Further, in lack of B_1 there is an accumulation of lactic and pyruvic acids in the blood and especially in the brain.‡ Convulsive symptoms in pigeons in extreme lack of B_1 are probably due to this; when vitamin B_1 is added to the brain of such pigeons *in vitro* the oxygen utilisation of the tissue is raised. It plays a part as a co-enzyme for pyruvate carboxylase (which co-enzyme is the pyrophosphate of thiamin), in the metabolism of carbohydrates (p. 94). It also stimulates conversion of carbohydrate into fat.§

Crystalline, vitamin B_1, was isolated by Windaus and by Otake; its formula was shown by Williams ‖ to be :—



Vitamin-B_1 (Hydrochloride of Thiamin.)

It has been synthesised,¶ and can be estimated by measuring the fluorescence of its oxidation product, thiochrome.

The international standard for vitamin B_1 is the pure substance. The international unit has been defined as 3 μg. of this. The requirement depends on the rate of metabolism. For average human individuals it is at least 300 units (1 mg. thiamin) daily. Up to 25 per cent. of ingested vitamin B_1 is excreted in the urine. Symptoms of lack include fatigue, loss of appetite, dyspnœa on exertion, muscle cramp, neuritis, and disturbances of sensation in the skin.

The richest sources of B_1 are dried yeast (10–20 I.U./g.), the germ of cereals, such as wheat and rice, peas, etc. White flour (72 per cent. extraction) contains very little, but the higher extraction flours (up to 85 per cent.) include the germ and so provide the vitamin (up to 1 I.U./g.).

The daily requirements for man are related to the calorific value of the (non-fat portion of the) diet, and can be calculated from Cowgill's formula :—

$$\text{Vitamin } B_1 \text{ intake (in International Units)} = \text{Calorific intake} \times \text{Wt. (Kg.)} \times 0{\cdot}0284.$$

---

* Webster and Armour. *Proc. Soc. Exp. Biol. Med.*, 1934, **31**, 463.

† Birch and Harris. *Bioch. J.*, 1934, **28**, 602.

‡ Kinnersley and Peters. *Bioch. J.*, 1930, **24**, 711.

§ McHenry and Gavin. *J. Biol. Chem.*, 1939, **128**, 45.

‖ Williams. *J. Amer. Chem. Soc.*, 1936, **58**, 1063.

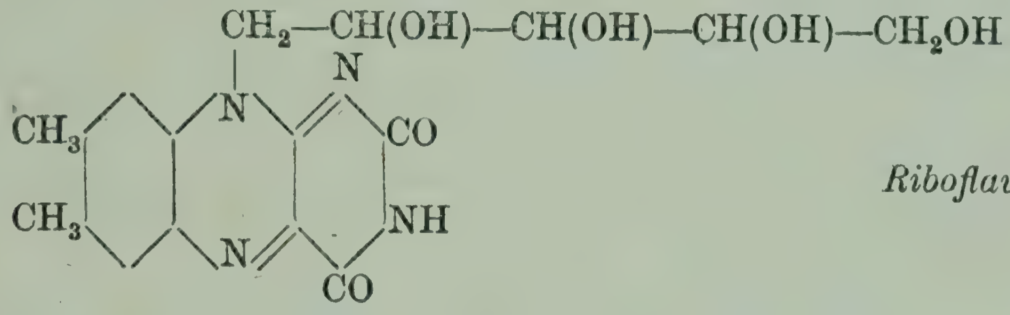
¶ Williams and Cline. *Ibid.*, p. 1504.

The minimum requirement of an adult on a diet of about 3,500 calories would be 550 units (about 1·7 mg.). Many modern diets supply only about 350. There is, however, considerable danger of vitamin B_1 deficiency in patients on restricted diets, or where individuals cut their own food intake down deliberately. The polyneuritis of chronic alcoholism is often traceable to shortage of vitamin B_1 consequent on reduced food intake. The beri-beri "danger-line" appears to be about 200 units (= 0·6 mg.) daily.

The optimal amount is about 500–1,000 units daily, according to the calorie intake. Children probably require about 400 units daily. The nutritional state of a subject regarding B_1 can be ascertained by estimation of the content of the vitamin in the urine before and after administration of a test dose.*

Thiamin provides an example of competition factors. If the sulphur in its molecule is replaced by a vinyl group, the substance pyrithiamine is obtained. When this is given to mice, symptoms of thiamin deficiency ensue; these are cured by increasing the thiamin to at least $\frac{1}{30}$th of the dose of pyrithiamine given.

Riboflavin (also known as lactoflavin). Its chemical name and structural formula are shown below :—



Riboflavin.

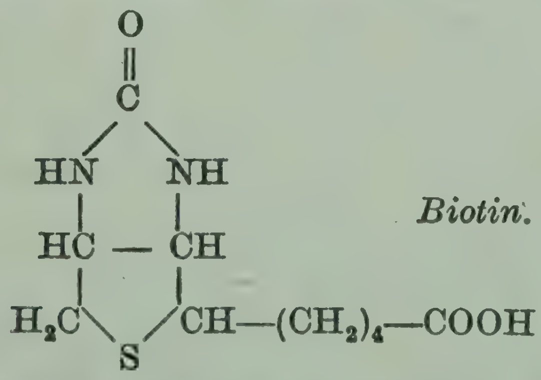
6.7. dimethyl-9-ribityl-isoalloxazin.

It is a yellow substance, present in milk whey, liver, kidney, heart, etc., which promotes growth, prevents cataract, keratitis and alopecia in rats, but does *not* prevent pellagra. Its deficiency in man causes lesions in the mouth (cheilosis) and tongue, vascularisation of the cornea, and neuritic symptoms. It is an important constituent of the "yellow respiration enzyme" of Warburg and occurs in the tissues in that form, viz., as a flavin-adenine-dinucleotide (*v.* p. 94), and hence plays an important part in oxidation processes. The daily requirement for man is 2–3 mg. It can be estimated by fluorescence, or by its absorption spectrum.

Symptoms of riboflavin deficiency can also be caused by administration of a phenazine analogue of riboflavin which acts as a competitor.

Pantothenic acid is a derivative of β-alanine and appears to be present in all living cells. It possibly plays a rôle in the intermediary metabolism of carbohydrates, and its absence from a diet leads to failure of growth and widespread and varied disturbances, including effects on the central nervous system.

Biotin (vitamin H) is a regular constituent of all normal human diets; it neutralises the effects resulting from a diet very rich in white of egg. Its formula is :—



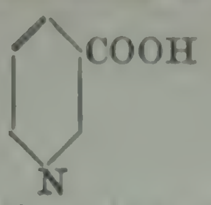
Biotin.

Lack of the vitamin has been experimentally produced in man, and results in a dermatitis of peculiar type. Egg white contains a substance ("avidin") capable of combining with biotin and thus depriving the body of it.

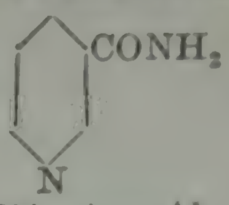
Nicotinamide (niacin or the P–P factor), is the vitamin which prevents pellagra, a disease characterised by dermatitis, diarrhœa and sometimes dementia, which is met with in peoples feeding mainly on maize. The disease called "black-tongue" in dogs is similar. The substance has been shown to be nicotinamide, and that substance is also formed in the body from nicotinic

---

* WANG and HARRIS. *B.M.J.*, Oct. 9th, 1943, p. 451.



Nicotinic Acid.



Nicotinamide.

acid, so that either will cure pellagra in man * or black-tongue in dogs.† It is widely distributed in most foodstuffs, particularly in meat, yeast, liver, fish, wheatmeal flour and beer, though maize contains but little. Average adults excrete 2–8 mg. daily in the urine, chiefly as the nicotinamide methochloride; if a dose, *e.g.* 100 mg., is given by mouth the excretion is raised promptly.‡

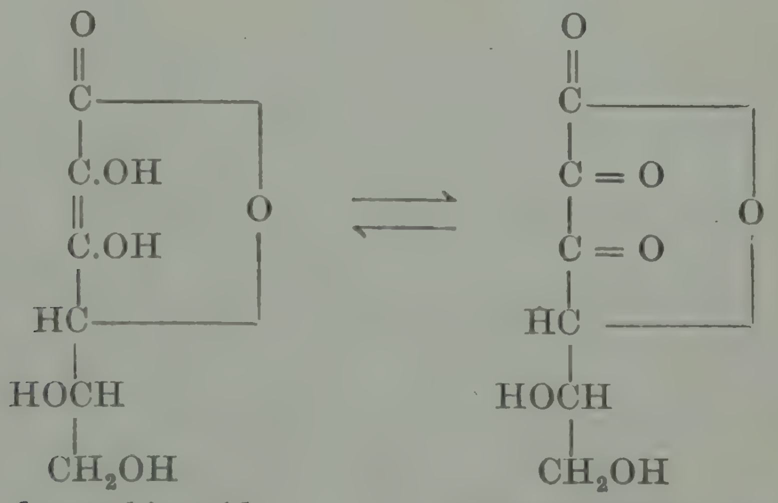
It is worthy of note that, once more we have a vitamin which functions as a co-enzyme. In this instance, nicotinic acid is part of the co-dehydrogenase complex and therefore probably plays a part in oxidative mechanisms (pp. 93, 98). Daily minimal requirement is about 10 mg.: intake in excess of this improves efficiency of strenuous work.

The analogues 3-acetyl pyridine and pyridine-3-sulphonic acid can, by competition, antagonise niacin. Maize appears to contain a similar pyridine compound capable of competing with niacin, and this probably explains why a diet of maize so readily causes pellagra unless the dietary niacin is augmented.

Choline has the effect of producing a lipotropic action (*v.* Ch. XLVI).

Vitamin B_{12} the hæmopoietic factor (*v.* p. 520).

Vitamin C is the antiscorbutic vitamin. Scurvy has been long known: it occurs when men are cut off for a long time from fresh food, especially vegetables. Its main symptoms are weakness, anæmia, skin eruptions, small hæmorrhages under the skin, impaired healing of wounds or fractures, liability to infection, hæmorrhage from the gums and mucous membranes (p. 639), and a liability to peptic ulceration. In children the long bones cease to grow, and there is tenderness of the legs. Captain Cook in 1776 § showed that it could be prevented or cured in sailors by the administration of malt or fresh vegetables. Oranges and lemons are specially rich in vitamin C. We now know ‖ that vitamin C is *l-ascorbic acid*, a hexuronic acid readily oxidisable to dehydro-*l*-ascorbic acid:—



l-ascorbic acid. *Dehydro-l-ascorbic acid.*

In the tissues it exists mainly in the reduced form. It was isolated from oranges, and also from the suprarenal cortex by Szent-Györgyi.¶ It is now made synthetically. It is present in small quantities in fresh meat and eggs, absent from most grains, but is formed when these grains germinate. When fresh

* Spies, Cooper and Blankenhorn. *J. A. M. A.*, 1938, **110**, 622.

† Goldberger, Joseph and Wheeler. *Publ. H. Rep.*, 1928, **43**, 172.

‡ Ellinger and Coulson. *Bioch. J.*, 1944, **38**, 265.

§ Cook, J. *Phil. Trans. Roy. Soc.*, 1776, **66**, 402.

‖ Svirbely and Szent-Györgyi, *Nature*, 1932, **129**, 576; *Bioch. J.*, 1932, **26**, 865; Waugh and King, *Journ. Biol. Chem.*, 1932, **97**, 325; Harris and Ray, *Bioch. J.*, 1933, **26**, 2067.

¶ Szent-Györgyi. *Bioch. Journ.*, 1928, **22**, 1387.

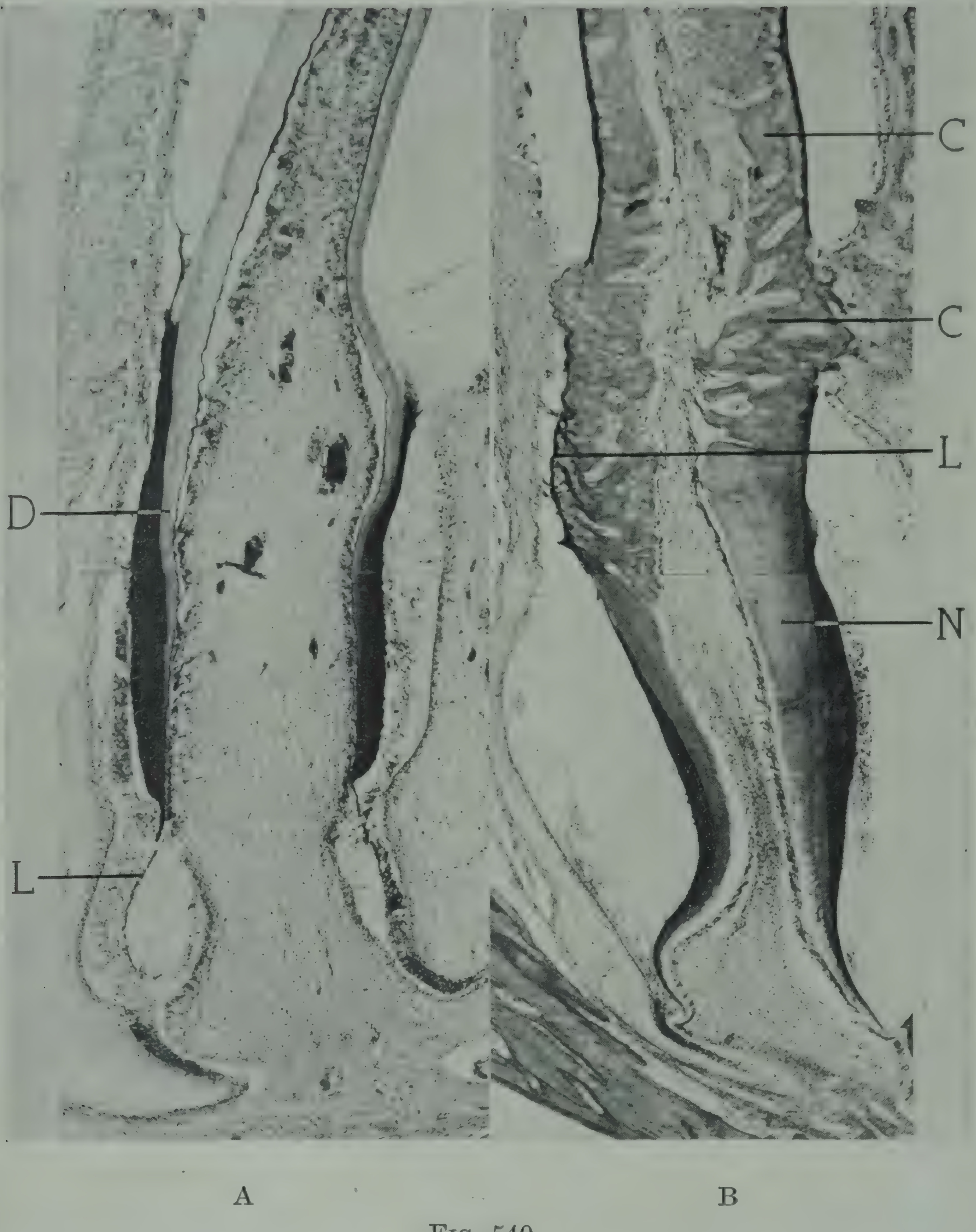


FIG. 540.

A. Growing end of tooth of guinea-pig deprived of vitamin C for fifteen days and then killed. Photomicrograph × 115.

D. Normal dentine formed before the deprivation of vitamin C.
L. Linear deposit of calcium salts deposited in place of dentine during attack.

B. Growing end of tooth of guinea-pig deprived of vitamin C for twenty days, then transferred to full diet for ten days and killed. Photomicrograph × 115.

N. Normal dentine formed since the attack.
C. Calcific scar formed since the attack.
L. Linear deposit of lime salts deposited in place of dentine during the attack of scurvy.

(Courtesy of DR. E. W. FISH.)

vegetable food cannot be obtained, scurvy may be prevented by feeding with beans which have been allowed to germinate, or by giving the pure ascorbic acid. In an acid medium this vitamin will stand exposure to boiling point for some time, but it is destroyed by oxidation almost at once on heating in an alkaline medium.

Vitamin C probably plays some part in the oxidation-reduction processes of the cell, acting as a hydrogen transporter. Blood serum, except in deficiency, contains about 0·8 mg./100 ml. : deficiency is indicated if below 0·4 mg./100 ml.[*] Its main actions in the body are twofold ; it is essential to the functioning of formative cells, such as fibroblasts ; and it is necessary for proper formation of intercellular collagenous substance. Hence a lack of vitamin C leads in guinea-pigs to arrest of formation of enamel and cementum of the teeth, and it has been shown [†] that there is a breakdown in the formation of the collagenous part of the matrix. This is well seen in the persistently growing teeth of the guinea-pig if the animal is deprived of vitamin C for, say, twenty days. The normal dentine produced before the attack appears in strong contrast to the irregular linear deposit of calcium salts (Fig. 540 (A)) held together by distorted traces of organic matter which is thrown down during the attack. If the animal is restored to a normal diet for ten days before being killed, fresh odontoblasts are seen to have developed and to have provided the tooth with a new normal growing end. At the same time a belated mass of calcified collagenous matrix is deposited amongst the old degenerated odontoblasts (Fig. 540 (B)).

VITAMIN CONTENT OF FOODS

Food.	A Int. Units per 100 gm. or 100 c.c.	B_1 mg. per 100 gm. or 100 c.c.	C mg. per 100 gm. or 100 c c.	D Int. Units per 100 gm. or 100 c.c.
Butter	2,000–5,000	0·03–0·3	—	10–100
Cheese (Cheddar)	5,500	0·1	—	—
Egg Yolk	8,800	0·3	0–1·3	100–250
Milk (cow's)	300–750	0·07	0·6–2·8	1·7–3·8
Bacon	30	0·3	Trace	—
Beef, lean	120	0·1–0·3	1·8	—
Liver	6,000–160,000	0·5	25–75	up to 50
Fish	up to 1,500	0·1–0·2	1·7	100–800
Fish Roe	5,000	0·25	4·0–14·0	250–990
Apples	70	0·125	0·1–20	—
Lemon Juice	—	0·05	25–70	—
Nuts	100	0·24–0·9	15–30	—
Tomatoes	4,280	0·12	12–40	—
Cabbage, etc.	900	0·07–0·25	20–124	—
Carrots	2,000	0·18	1·8–31	—
Peas	800–1,300	0·12–0·3	4·8–40	—
Potatoes	35–70	0·1 -0·2	10–36	—
Spinach	12,230	0·06–0·21	6–124	—
Bread, white	—	0·036–0·1	—	—
Bread, wholemeal	—	0·16–0·25	—	—
Oatmeal	0–20	1·0	—	—
Rice, whole	50	0·06–0·27	—	—
Rice, polished	—	—	—	—
Yeast	—	1–12	—	—

---

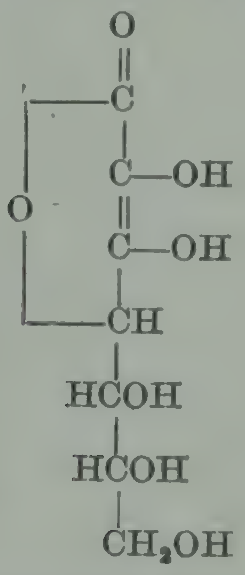
[*] PRUNTY and VASS. *Bioch. J.*, 1943, **37**, 623.

[†] FISH and HARRIS. *Phil. Trans.*, 1933, **223B**, 489.

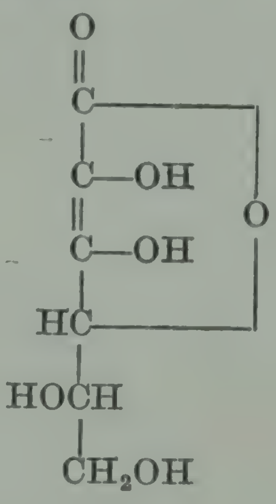
Growing scorbutic animals also show impaired bone formation, both as regards matrix and calcification; and the phosphatase content in costochondral junctions is reduced. It has been suggested that ascorbic acid is essential for the formation of the phosphatase-impregnated matrix in which calcification can occur.*

For experimental work on vitamin C, it is usual to employ guinea-pigs. The international unit is 0·05 mg. pure *l*-ascorbic acid, of which lemon juice contains about 10 units per ml. A young guinea-pig requires about 0·7 mg. per day to prevent scurvy.† The fœtus receives its store of ascorbic acid from the mother. Children and adults require a maximum of about 1,000 units (50 mg.) of ascorbic acid daily (minimum 25 mg.). It is best supplied by fruits, *e.g.* lemons, oranges and black currants, or vegetables. The amount of vitamin C in materials can be determined chemically by titration with "redox" dye (2 : 6 dichlorophenolindophenol).‡ By this means it can be shown that normally most of the ingested ascorbic acid is excreted in the urine, the daily amount being 1 to 2 mg. in children and up to 30 mg. in adults. Scorbutic children excrete very much less than normal amounts (less than 15 mg. daily in adults indicates deficiency). If a large dose (100 mg.) of vitamin C be given to a normally nourished individual it is rapidly excreted, whereas when given to scorbutics the output is scarcely altered, showing that the substance is not excreted until a certain level of saturation has been reached.§ The vitamin may also be estimated by a ferricyanide reduction method,‖ and spectrometrically.

The analogous compound glucoascorbic acid, when given in the diet competes with ascorbic acid, and causes scurvy, even in mice, which do not normally need ascorbic acid: the condition can be cured by suitable doses of ascorbic acid.



Glucoascorbic acid.



Ascorbic acid.

Folic acid is a name given to a group of conjugated derivatives of pteroylglutamic acid; the latter substance, now synthetically made, is also sometimes called folic acid. Pteroylglutamic acid or its derivatives are widely dispersed in animal and plant cells, and can exert curative effects in pernicious anæmia (p. 520) though the nature of these effects is not clear.

Vitamin P is found together with ascorbic acid, is heat stable, and its lack is largely responsible for the capillary hæmorrhages which occur in scurvy. It is also called citrin.

"*Citrin*," found in the rind and juice of citrous fruits, is a mixture of two flavone pigments, viz. a rhamnose-glucose glycoside of hesperidin, and a glycoside of eriodictyol. The human requirement is stated to be 50 mg. daily.

Vitamin Content of Foodstuffs. The vitamin content of some common foodstuffs is shown in the Table on p. 829.

For Reference

Almquist (1941). Vitamin K. *Physiol. Rev.*, **21**, 194.
Bills (1935). Physiology of the Sterols, Including Vitamin D. *Physiol. Rev.*, **15**, 1;
Bourne (1936). The Rôle of Vitamin C. *Physiol. Rev.*, **16**, 442.

---

* Bourne. *J. Physiol.*, 1943, **102**, 319.
† Lund, Spur and Fridericia. *Bioch. Journ.*, 1934, **28**, 1825.
‡ Harris and Ray. *Bioch. J.*, 1933, **27**, 303.
§ Harris and Ray. *Lancet*, 1935, i., 71.
‖ Tauber and Kleiner. *J. Biol. Chem.*, 1935, **108**, 563.

Clausen (1934). The Influence of Nutrition upon Resistance to Infection. *Physiol. Rev.*, **14,** 309.

Clausen (1942). The Absorption of Vitamin A and its Storage in the Tissues. *Harvey Lects.*, p. 199.

Coward (1938). The Biological Standardisation of the Vitamins. (London : Baillière, Tindall and Cox.)

Dann and Darby (1945). The Appraisal of Nutritional Status (Nutriture) in Humans. *Physiol Rev.*, **25,** 326.

Doisy *et al.* (1941). Vitamin K. *Chem. Rev.*, **28,** 477.

Drill (1943). Inter-relations between Thyroid Function and Vitamin Metabolism. *Physiol. Rev.*, **23,** 355.

Eddy and Dalldorf (1941). The Avitaminoses. (London : Baillière, Tindall and Cox.)

Elvehjem (1940). Relation of Nicotinic Acid to Pellagra. *Physiol. Rev.*, **20,** 249.

Hertz (1946). Biotin and the Avidin-Biotin Complex. *Physiol. Rev.*, **26,** 479.

Jukes and Stokstad (1948). Pteroylglutamic Acid and Related Compounds. *Physiol. Rev.*, **28,** 51.

King (1936). Vitamin C, Ascorbic Acid. *Physiol. Rev.*, **16,** 238.

M.R.C. Report (1949). Vitamin A Requirements of Human Adults. *H.M. Stationery Office.* Special Reports No. 264.

Pappenheimer (1943). Muscular Disorders associated with Deficiency of Vitamin E. *Physiol Rev.*, **23,** 37.

Reid (1943). Inter-relations of Calcium and Ascorbic Acid, etc. *Physiol. Rev.*, **23,** 76.

Snell (1948). Use of Micro-organisms for Assay of Vitamins. *Physiol. Rev.*, **28,** 255.

Spies (1947). Experiences with Folic Acid. (Year Book Publ. Inc.)

Wolbach and Bessey (1942). Tissue Changes in Vitamin Deficiencies. *Physiol. Rev.*, **22,** 233.

Woolley (1947). Recent Advances in the Study of Biological Competition between Structurally related Compounds. *Physiol. Rev.*, **27,** 308.

CHAPTER XLIII

THE NORMAL DIET OF MAN

Examination of the actual requirements of the normal human individual involves a consideration not only of the total energy requirements of man according to age, size, sex and occupation, but also of the proper distribution of the foodstuffs in the diet.

We have seen that the energy requirements are a function of the body surface. The average height and weight of English adults is 171 cm. and 70·3 kilos., corresponding to a surface of 1·8 square metres (Fig. 527, p. 794). At 40 Calories per square metre per hour the average man will thus have an hourly basal metabolism of about 72 Calories. When he is taking food this amount must be increased by about 10 per cent., and if he is up and about a further increase in metabolism will be evoked by the muscular movements, exposure to cold, etc. If we divide the twenty-four hours into three portions—eight hours' sleep, sixteen hours awake—(eight hours' work)—we must allow the basal metabolism for the eight hours' sleep, basal metabolism increased by about 30 per cent. for the eight hours during which the man is up and about and taking meals, and the basal metabolism plus an additional amount for work during the eight working hours. It is by no means easy to decide what figure should be accepted as representing the average work of an adult man. It is usual to take this as involving an energy output of 1,000 Calories. Of these Calories 20 per cent., or even more, may under favourable conditions be transmuted into mechanical work, so that 1,000 Calories should be sufficient to provide for the performance of about 100,000 kilogrammetres of external work. With these assumptions we may suggest the following energy requirements for an average adult working man :

8 hours' sleep at 71 (Basal Metabolism)	568	Calories.
8 hours awake at 92 (Basal Metabolism plus 30 per cent.)	736	,,
8 hours' work (Basal Metabolism plus 1,000 Calories)	1,568	,,
Total	2,872	,,

Some additional expenditure of energy will be involved in travelling, so that we are justified in accepting the figure usually given for the daily energy output of an average working man as 3,000 Calories. For purposes of computation, about 400 Calories are usually added to allow for loss in preparation and in digestion, so that for the average man about 3,400 Calories must be purchased.

A man in sedentary employment—*e.g.* typewriting, clerking, etc.—may not expend more than 200 to 400 Calories in the performance of external work ; mental exertion causes no appreciable rise in the metabolic exchanges. On the other hand, a heavy worker, such as a navvy, may easily expend 2,000 or more Calories a day in the performance of external work.

The energy requirements of women will show corresponding variation according to occupation, but will be in general less in consequence of their smaller size and surface. Du Bois has shown that women have a lower basal metabolism per square metre of surface, viz. 37 Calories per hour as against 40 in the case of men. The average height and weight of Englishwomen are 159·3 cm. and 55·7 kilos., corresponding to a surface of 1·511 square metres and to an hourly output of 56 Calories. In estimating the total energy requirements of an average woman, we may reckon her work as two-thirds that of a man, *i.e.* as 600 Calories per day. We then arrive at the following Table for the energy requirements of the average woman :

8 hours' sleep at 56	448	Calories.
8 hours awake at 73 (Basal Metabolism plus 30 per cent.)	584	,,
8 hours' work (Basal Metabolism plus 660 Cals.)	1,108	,,
Total	2,140	,,

We shall not be far wrong, then, in accepting the usual estimate of a woman's net requirements as four-fifths of those of a man, viz. 2,400 Calories, or the gross value as purchased as 2,800 Calories.

In children a computation of the energy requirements is rendered more difficult by two factors. In the first place the metabolism is more energetic than in the adult, even if we take into account the relative surfaces in the two cases. Thus the basal metabolism per square metre in boys of varying ages was found by Du Bois to be as follows :

Mean age.	Basal metabolism, Cals. per square metre per hour.
6·5	57·5
12·6	50·4
13·7	49·4
16·5	43
19·25	40·7

In addition to this increased basal metabolism, if the processes of growth are to occur normally, there must be an excess of matter taken in over that excreted, since the growing child is always putting on weight. Between the ages of eleven and sixteen both sexes put on weight at about 4 kilos. a year, which is equivalent to adding to the body a store of about 800 Calories per month. The second difficulty is due to the fact that children are in a constant state of muscular activity, which is of great value in procuring healthy nervous and muscular development. The degree of this activity seems related to the amount of food taken—the increased energy of a child directly after a meal is a familiar phenomenon. If food is withheld or diminished this activity declines, but in this way energy is saved and the whole of the lessened food may go to maintain the temperature of the body and the normal processes of growth.

We may thus obtain a development of the child not differing appreciably from the normal on widely varying amounts of food, and it becomes difficult to say what is the average or the optimum amount that a child should have.

The table below shows the requirements of the different sexes and ages in terms of the "average man" taken as 1.

MAN-VALUES AT AGES (CATHCART AND MURRAY) (B.M.A. Report, 1933)

Ages.	Man-value.	Calories.
Adult male	1·00	3400
,, female	0·83	2840
Child 1 and under 2 years	0·30	1020
,, 2 ,, 3 ,,	0·40	1360
,, 3 ,, 6 ,,	0·50	1700
,, 6 ,, 8 ,,	0·60	2040
,, 8 ,, 10 ,,	0·70	2380
,, 10 ,, 12 ,,	0·80	2720
,, 12 ,, 14 ,,	0·90	3000
Old person (65 +)	0·75	2550

The figures in the last column represent the Calories as purchased for the average individual, and are therefore equivalent to the food which must be digested and absorbed plus any loss. No food undergoes complete digestion, the average loss being about 10 per cent. with an ordinary mixed diet, as is shown in the following Table :

Percentage of Foodstuffs Absorbed.

	Protein.	Fat.	Carbohydrate.	Ash.	Total Energy.
Average of 5 experiments	92·6	94	97·7	77·4	90·5

A further allowance must be made for unavoidable waste in cooking, etc.

Three thousand net Calories (*i.e. energy* requirements) require therefore the purchase of food with a Calorie value of 3,400.

If the food consists mainly of vegetable products it may be necessary to increase still further the allowance for loss in digestion, not more than 80 per cent. of the protein being absorbed as a rule.

These values are for average well-nourished individuals. It has already been pointed out that in a state of semi-starvation, accompanied by loss of weight, the basal requirements may be reduced considerably. As a result of a number of experiments Chittenden * came to the conclusion that there was a definite advantage to be gained by reducing not only the protein intake of the body but also the total food. But the experience of Germany during the last three years of the 1914 war showed that a forced limitation of food to about two-thirds of what we have put down as normal, resulted finally in a decrease of efficiency of mind and body and in a marked diminution in the resistance to infection, especially to tuberculosis.

Distribution of Foodstuffs in a Normal Diet. Before 1939 there was fairly general agreement that the average diet of an adult man should be somewhat as follows :—

Daily Requirements of an Adult Male of Calories, Protein, Fat, and Carbohydrates (B.M.A. Report, 1933)

Total Calories required : 3400 in food as purchased

Derived from—	Quantity. g.	Percentage of total Calories.	Number of Calories.
First-class protein .	50	6	205
Second-class protein .	50	6	205
Fat	100	27	930
Carbohydrate . .	500	61	2050
Totals . . .	700	100	3390

Since the war nutritional requirements have been more parsimoniously scrutinised, and revised in the light of experience, or the twilight of necessity. A more recent recommendation is that by the Food and Nutrition Board, National Research Council, U.S.A., which was adopted as a goal in 1945.

Quoted from the Recommendations of the Food and Nutrition Board, N.R.C., U.S.A.

	Calories.	Protein. g.	Ca. g.	Fe. mg.	Vit. A. int. Units.	Vit. B$_1$. mg.	Ribo-flavin. mg.	Niacine. mg.	Ascorbic Acid. mg.	Vit. D. int. Units.
Man (70 kg.),										
Sedentary .	2500	70	0·8	12	5000	1·2	1·6	12	75	
Very active .	4500					2·0	2·6	20	—	
Woman (56 kg.),										
Sedentary .	2100	60	0·8	12	5000	1·1	1·5	11	70	
Pregnant (latter half) . .	2500	85	1·5	15	6000	1·8	2·5	18	100	400–800
Nursing . .	3000	100	2·0	15	8000	2·0	3·0	20	150	400–800
Children, 1–3 .	1200	40	1·0	7	2000	0·6	0·9	6	35	400

Wide variations are possible in the relative proportions of the foodstuffs without injury to health. It is important, however, to consider the limits within which these variations are permissible.

Proteins. From 10 to 15 per cent. of the total Calorie requirements should consist of protein, *i.e.* an average mixed diet with a Calorie value of 3,400 will contain about 10 grammes of protein. Where animal protein is entirely absent from the food, a larg amount of vegetable proteins will be necessary, since they are of lower " biological value " and it is safe to lay down as a rule that for an average man the protein ration should not b diminished below 1 g./kg., or 70 g. per day. This should include different kinds of protei and at least half of it should be first-class, or animal protein. For infants of 1–3 years a much as 4 g./kg. may be needed : this requirement is lessened with advancing age, and a

---

* Chittenden. " Physiological Economy in Nutrition." 1904 (Stokes) ; " Th Nutrition of Man," 1907.

17–18 years, is about 2 g./kg. In recent wars and famines large populations have had to submit to lowered protein intake, *e.g.* in the siege of Madrid, total protein 66g. (animal 4·8 g.); in Great Britain in 1944 total protein 70 g. (animal 33 g.) daily, with no evident ill-effects.

Butcher's meat contains from 15 to 20 per cent. protein, and from 15 to 30 per cent. fat. Owing to the high specific dynamic action of protein, a protein meal raises the production of heat in the body, which is a distinct disadvantage when work has to be accomplished at a high external temperature. On the other hand, a diet in which there is a large proportion of meat is of value to men in sedentary occupations in a cold or temperate climate. The idea that the heavy worker requires a large supply of meat is of doubtful foundation, though in occupations involving exposure to cold and wet it will add to the comfort of the individual by keeping him warm.

Meat is not only easily presented in a palatable form, but its flavour renders other kinds of foods acceptable. It is thus habit rather than strict physiological principles which will probably govern the quantity of meat regarded as desirable.

Carbohydrates are as a rule the most abundant and cheapest constituents of any diet, and form the greater part of most diets. A certain amount of carbohydrate is essential for man. In the absence of carbohydrate in the food there is a deficient oxidation of the fats, giving rise to acid intoxication.

Fats can be produced by the animal body from carbohydrate, and part of either the fat or the carbohydrate of a normal diet is mutually interchangeable for an isodynamic amount of the other. Although some men can remain healthy on almost fat-free diets provided vitamins are available, some fat is essential, and quantities greater than the minimum amount are desirable. In practice we have to provide for each race a minimum desirable quantity of fat, which will differ in different races and be larger in northern than in southern races. Thus the Japanese soldier is content with 20 grammes of fat daily, while in our country the average man would not be content with less than 75 grammes of fat per day. Seventy-five grammes of fat has a Calorie value of about 25 per cent. of the daily requirements, and we may conclude as a general rule that one-fourth of the total energy of the diet should be supplied in the form of fat.

In recent times of deprivation due to war, the shortage of fats ("guns before butter") was the most hard to bear. There are probably three reasons for this craving for fat. In the first place, whereas the greater part of the carbohydrate is absorbed within three hours after food has been eaten, the chief absorption of fat occurs between five and six hours after a meal. On this account individuals deprived of fat get hungry some time before the next meal. The lumbermen of Canada satisfy their huge needs in Calories—6,000 to 8,000 a day—by a diet in which the fat represents 35 to 40 per cent. of the Calories.

In the second place, weight for weight, fat has more than double the Calorie value of carbohydrate. But the difference in bulk is still greater, since most other foods contain a considerable proportion of water.

In the third place, carbohydrates are more subject to fermentative changes in the intestines than fats. We may conclude then that although man can for a time dispense with fat, yet, to develop his full efficiency as a working machine, fat is an essential ingredient of his diet.

We thus see that a dietary to maintain a sedentary man in health must fulfil the following requirements:

(1) It must be palatable. Appetite is an essential factor for the digestion and assimilation of the food, so that good cooking and appropriate flavouring becomes an important condition for the maintenance of health.

(2) It must have a sufficient Calorie value (3,400 for the average man).

(3) It must contain protein, fats and carbohydrates. The quantity of each of the first two should be not less than 70 and preferably not less than 100 grammes a day. Of the protein, there should be at least 50 grammes of first-class (*i.e.* animal) protein daily.

(4) It should include a certain proportion of fresh foods, such as green vegetables, meat and eggs, and, in the case of children, milk, in order to supply the necessary vitamins.

(5) It must contain a proper proportion of the salts, especially sodium, calcium, potassium, chlorides and phosphates.

Actual Diets. Statistical inquiries into the food consumption of controlled groups of individuals, *e.g.* soldiers, prisoners, school children,* families, or even of whole nations have given results which lend further support to those obtained from laboratory observations. Thus the population of the United Kingdom in 1911 was about 45¼ million persons: its man-value, as calculated from the age-distribution, had an average of 0·835, so that the total man-value was about

---

* FRIEND. "The Schoolboy" (Heffer, Cambridge), 1935.

37¾ million. The Calorie value of all the food grown or imported during 1911 was about 47 billion Calories, which works out at about 1¾ million Calories per man-value per annum or roughly 3,400 Calories per day.

As an indication of the great improvement in standards in modern times as contrasted with that two centuries ago, the investigations of Friend, Chief Medical Officer of Christ's Hospital School, may be quoted. The requirements of boys of school age are about 2,800 Calories. The diets in the years 1704 and 1929 are given below :

Diet of Schoolboys at Christ's Hospital

(After G. E. Friend)

	In 1704	In 1929
Protein	39 grammes	93 grammes
Carbohydrate	188 ,,	394 ,,
Fat	26 ,,	100 ,,
Calories as beer	730	*Nil*
Total Calories	1,900	2,930

The study of the dietaries of families of different incomes * yielded similar results and showed in particular that there was a great dietary disparity, both qualitative and quantitative, between the different classes of the community. It has been claimed that before World War I, up to 30 per cent. of the population, owing to a combination of poverty and ignorance were undernourished, and a still larger proportion suffered from qualitative defects of diet. The following Table from Rowntree has often been quoted to illustrate the relative Calorie deficiency in the poorest classes as compared with the contemporary standards of diet.

Dietaries of Various Classes in 1900

(After Rowntree †)

		Protein.	Calories, daily
In York	I. Wages below 26*s.* weekly	70	2,574
	II. Wages over 26*s.* weekly	96	3,390
	III. Servant-keeping classes	105	3,807
Agricultural Labourers, Midlands		74	2,868
Able-bodied paupers, York Workhouse		156	3,702
Convicts, hard labour		177	4,159

Since that war more thought has been taken to ensure that the poorest classes do not fall below the border-line of minimum requirements, though it cannot be doubted that prior to the late war there were still many who, for various reasons, failed to get proper nourishment. The position in 1949 is that the disparity between such classes as can still be distinguished has practically vanished ; the uniform diet is apparently now at about the level of Rowntree's poorest class : this is now thought to be adequate.

Rationing of food began in Great Britain in January, 1940 (butter, bacon, sugar) ; then followed meat (March, 1940), the National Milk Scheme in July, 1940, eggs in June, 1941, the points system in December, 1941. During the 1939 war the policy in this country ‡ was, roughly, to have ample supply of carbohydrate (potatoes, bread) which ensures adequate calories, and to distribute the sources of fat, and of protein food, such as meat and cheese, more equally. Protein supplies were also augmented by Lend-Lease : in 1940 it was down to 80 g. (36 g. animal) daily. Control of milk, with additional supplies to pregnant and nursing mothers, and to infants ; supply of vitamins A and D to pregnant women and to infants ; rationing of butter, with fortification of margarine by the addition of 450 I.U. of vitamin A and 30 I.U. of vitamin D per oz., and increased consumption of carrots (for A), have ensured adequate vitamins A and D. Increased use of green vegetables, and, for children, black-currant syrup, supplied adequate vitamin C. The total fat, visible and invisible, was about 131g./day per head before the war: by the middle of 1941 it had fallen to 104 g., but by the end of 1943 recovered to 111 g.* The calorie deficiency was made up on carbohydrates (*e.g.* bread and potatoes) which were then unrationed.

---

* Cathcart and Murray. An Inquiry into the Diet of 154 Families of St. Andrews. M.R.C. Report No. 151, 1931, and 165, 1932.

† Rowntree. " Poverty, A Study of Town Life " (cited by Starling, " Feeding Nations " ; and by Drummond, " Lane Lectures " (1934)).

‡ Pyke. *Brit. Med. Bull.*, 1944, **2**, 228.

The bread question is interesting. Before the war, wheat flour was 72 per cent. extraction, *i.e.* only 72 per cent. of the grain was used in flour. The wheat germ, etc., which was rejected, contained more than half the vitamin B_1 of the grain. At the end of 1941 it was decided to extract to 85 per cent., and the resulting National wheatmeal contained 1·3 mg. of B_1 per lb., as compared with 0·5 mg. in white flour. The germ included in the new flour also contained as much iron as the rest of the grain, but it also contained most of the phytic acid, in consequence of which the absorption of calcium was reduced.† As calcium was one of the most general deficiencies of our pre-war diet, this was a serious drawback. It was rectified by adding about 0·15 per cent. of calcium carbonate to the new flour, so adding about 180 mg. to the daily Ca intake.

After the end of the war the situation in Europe generally worsened. Bread rationing was begun in Great Britain on July 21st, 1946, and this greatly reduced the flexibility of the national diet; the subsequent rationing of potatoes in November, 1947, narrowed the safety margin still further, since, under the rationing system about half the total calorie supply was made up by bread, flour and potatoes. The following table shows the position in November, 1946, and includes bread, but not potatoes, in the ration; it should be compared with the U.S. goal, p. 834.

NUTRIENTS PROVIDED DAILY BY RATIONS, PRIORITY ALLOWANCES AND EXPECTED DISTRIBUTIONS OF MILK AND FRESH EGGS (NOV., 1946) IN GREAT BRITAIN ‡

	Calories.	Protein. g.	Ca. g.	Fe. mg.	Vitamin A. I.U.	B_1 mg.	Ascorbic acid. mg.
Normal male adult . .	1842	56	0·6	11	1398	0·77	3
Woman, pregnant . .	2440	85	1·5	13	5892	1·1	49
Woman, nursing . .	2085	69	1·1	11	1684	0·92	8
Children, 2–3 . . .	1745	57	1·1	8	5822	0·74	33

It will be seen that there is, except in the case of young children, a considerable gap in calories, while there is shortage of protein for adult males and nursing mothers; potatoes would normally help to fill this gap, which would otherwise only be filled by restaurant meals.

In these days of widespread semi-education it should be possible to instruct the masses in the matter of economical purchasing and correct cooking of proper food. As an example of pre-war diets which combine cheapness with proper nutritive value we may quote again from the B.M.A. Report. The first is the simplest type of diet for a man, and purchased in the most economical way would cost (in 1933) about 5*s*. per week. It provides no variety however.

To take another example, suppose a family consisting of man, wife, and child aged four. Its man-value, as deduced from the Table on p. 834, would be $1{\cdot}0 + 0{\cdot}83 + 0{\cdot}50 = 2{\cdot}33$, and an example of constituents for a diet costing (in 1933) about 14s. per week for the family is given on p. 838.

The conversion of these raw materials into adequate meals is a matter of cookery, but teachers of domestic science have translated these findings into actual cookery recipes enabling a family, *e.g.* of man-value about four, to obtain wholesome and palatable diet at a cost of about 23s. per week (in 1933).§

---

* PYKE. *Brit. Med. Bull.*, 1944, **2**, 228.

† MCCANCE and WIDDOWSON. *J. Physiol.*, 1942, **101**, 304.

‡ BRANSBY and MAGEE. *B.M.J.*, 1947, i, 525.

§ "Family Meals and Catering." (B.M.A., 1934. Price 6*d*.)

B.M.A. Diet No. 1. Bare Ration. No Variety Man-value 1

Item.	Quantity.	Protein (grams).	Fat (grams).	Carbohydrate (grams).	Calories.
Corned beef	1 lb	119·2	84·8	—	1278
Cheese	2 ,,	233·2	317·6	28·2	4022
Margarine	¾ ,,	0·7	288·5	—	2684
Flour *or*	7 ,,	320·6	51·1	2397·5	11,620
Bread	11¼ ,,	(367·9)	(10·1)	(2454·8)	(11,666)
Sugar	1¾ ,,	—	—	793·8	3255
Potatoes	3½ ,,	30·1	0·4	287·4	1306
Tea	¼ ,,	—	—	—	—
Fresh fruit and green vegetables	—	—	—	—	100
Total weekly quantities	—	703·8	742·4	3506·9	24,265
Daily quantities per man	—	100·5	106·1	501·0	3466

Total first-class protein, 352·4 grammes.
Daily first-class protein, 50·3 grammes.

B.M.A. Diet No. 11. Man, Wife and One Child, aged Three to Six Years. Man-value 2·33

Item.	Quantity.	Protein (grams).	Fat (grams).	Carbohydrate (grams).	Calories.
Beef	½ lb.	42·7	41·7	—	563
Minced meat	1 ,,	85·3	83·5	—	1126
Bacon	1 ,,	46·7	245·8	—	2478
Corned beef	1 ,,	119·3	84·8	—	1278
Liver	½ ,,	45·2	7·3	10·0	294
Cheese	2 ,,	233·2	317·6	28·2	4022
Milk	10½ pts.	196·8	214·2	285·6	3969
Fish (kipper)	¾ lb.	48·0	37·8	—	548
Butter	½ ,,	0·5	188·3	—	1752
Lard (dripping)	½ ,,	—	226·8	—	2110
Flour *or*	13½ ,,	618·3	98·6	4623·8	22,410
Bread	21½ ,,	(703·1)	(19·4)	(4691·3)	(22,296)
Sugar	2½ ,,	—	—	1134·0	4650
Jam	1 ,,	1·4	—	314·8	1296
Potatoes	7 ,,	60·2	0·7	574·7	2611
Peas (dried)	¾ ,,	69·4	2·0	194·1	1100
Yeast	¼ ,,	—	—	—	—
Tea	½ ,,	—	—	—	—
Oatmeal	1¼ ,,	67·5	48·8	396·9	2358
Rice	½ ,,	13·4	0·9	182·1	810
Treacle	½ ,,	3·7	—	135·9	572
Beans (butter)	½ ,,	42·2	1·6	141·1	766
Barley	¼ ,,	7·9	0·9	90·6	412
Fresh fruit and green vegetables	—	—	—	—	300
Total weekly quantities	—	1701·7	1601·3	8111·8	55,425
Daily quantities per man	—	104·3	98·2	497·4	3398

Total first-class protein, 817·7 grammes.
Daily first-class protein per man, 50·1 grammes.

During World War II, the Food Ministry in Great Britain gave, to those who heeded it, much useful information regarding diet [*]; though it is probable that this advice was least heeded where it was most needed. There is, however, no indication that the health of the population of Great Britain has suffered under rationing; vitamins A and C are the most likely ones to be lacking, but experimental supplementing with these and other vitamins showed no significant effects on the average, though in individual cases there were indications of some slight deficiency. A diet survey of 849 families in 1941 in eight localities in England and Scotland showed, however, that in groups where food expenditure was less than 9*s.* per head weekly, the calorie intake was inadequate.[†]

Intake of Nutrients per Head Daily (1941) (Bransby)

Food Expenditure per Head Weekly.	Prot. (g.)	Fat. (g.)	Carb. (g).	Cals.	Ca. (mg.)	Iron. (mg.)	Vit. A. (I.U.)	Vit. B_1. (I.U.)	Ascorbic Acid (mg.)
1. Under 5*s.* . .	57	55	270	1803	430	8	2062	214	15
2. 5*s.*–7*s.* . .	63	63	287	1967	493	9	1828	237	20
3. 7*s.*–9*s.* . .	68	75	292	2115	556	10	2620	259	23
4. 9*s.*–11*s.* . .	77	87	319	2367	634	12	2950	277	30
5. 11*s.*–13*s.* . .	80	88	341	2476	701	13	3353	286	34
6. 13*s.* and over .	92	105	360	2733	875	14	3891	313	40

For Reference

Amer. Med. Assoc. (1943). Handbook of Nutrition. (A.M.A., Chicago.)
British Medical Association (1933). Report of Committee on Nutrition. (Supplement to *B.M.J.*, November 25th.)
Cruickshank (1946). Food and Nutrition. (Edinb.: Livingstone.)
Drummond and Wilbraham (1939). The Englishman's Food. (London: Cape.)
Ershoff (1948). Conditioning Factors in Nutritional Disease. *Physiol. Rev.*, **28**, 107.
Marrack (1942). Food and Planning. (London: Gollancz.)
Sherman (1932). Chemistry of Food and Nutrition. (Macmillan.)
Starling (1919). The Feeding of Nations. (London: Longmans Green.)
Sydenstricker (1943). Nutrition under Wartime Conditions. (*Harvey Lects.* Lancaster Pa. Science Press Printing Co.)

Alcohol. When alcohol is taken by man in moderate quantities, up to 10 c.c. per hour, the greater part of it undergoes oxidation, mainly in the liver,[‡] *viâ* acetic acid, to carbon dioxide and water. About 2 per cent. of it, which escapes oxidation, is excreted unaltered by the lungs and kidney. This oxidation of alcohol is a result of true utilisation and alcohol can, in this way, act as a food and supply up to 40 per cent. of the daily calories, provided it is taken in small doses, often repeated, so as not to result in there being a high alcohol content in the blood.[§]

The chief effects of alcohol are on the nervous system. The latent period of responses is increased and the intensity of conditioned reflexes is reduced, while the normal balance between excitatory and inhibitory states is upset.[‖][¶] Hunger contractions are inhibited, and alcohol in moderate doses either does not alter, or else retards, the rate of emptying of the stomach.[**] Gastric secretion is stimulated by moderate doses, probably by release of histamine: concentrations over 10–15 per cent. increase the mucus secretion, and if frequently taken cause acute gastritis. Its value in a diet is chiefly that of an accessory or adjuvant, in exciting appetite by its taste and smell. When taken in large doses, this function is overshadowed by the poisonous action of the substance. The

---

[*] Pyke (1945). Manual of Nutrition. M.O.F. (H.M. Stationery Office). 1*s.*

[†] Bransby. *B.M.J.*, 1946, i, 832.

[‡] Lundsgaard. *C. R. Lab. Carlsberg*, 1937, **22**, 333.

[§] M.R.C. Special Report, No. 31.

[‖] Andreyev. *Arch. Internat. de Pharm. et de Thérap.*, 1934, **48**, 117.

[¶] Horsley Gantt. *Bull. Johns Hopk. Hosp.*, 1935, **56**, 61.

[**] Beazell and Ivy. *Q. J. Stud. on Alcohol* 1940, **1**, 45.

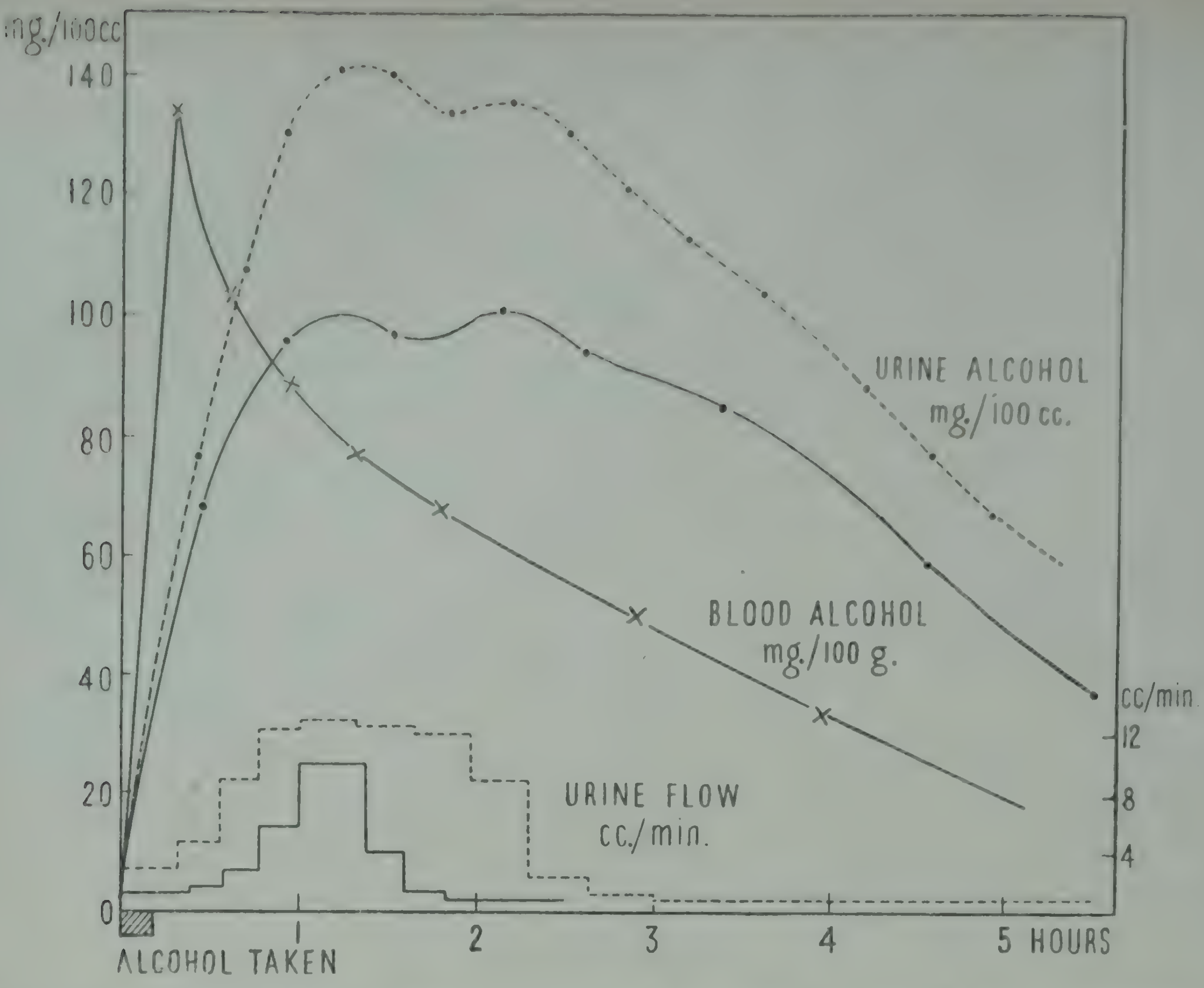


FIG. 541. Comparison of a naturally fast absorber of alcohol with a naturally slow one. The alcohol was taken on an empty stomach :—

×————× 0·76 g/kg.
·————· 0·83 g/kg.

(Courtesy of DR. M. GRACE EGGLETON.)

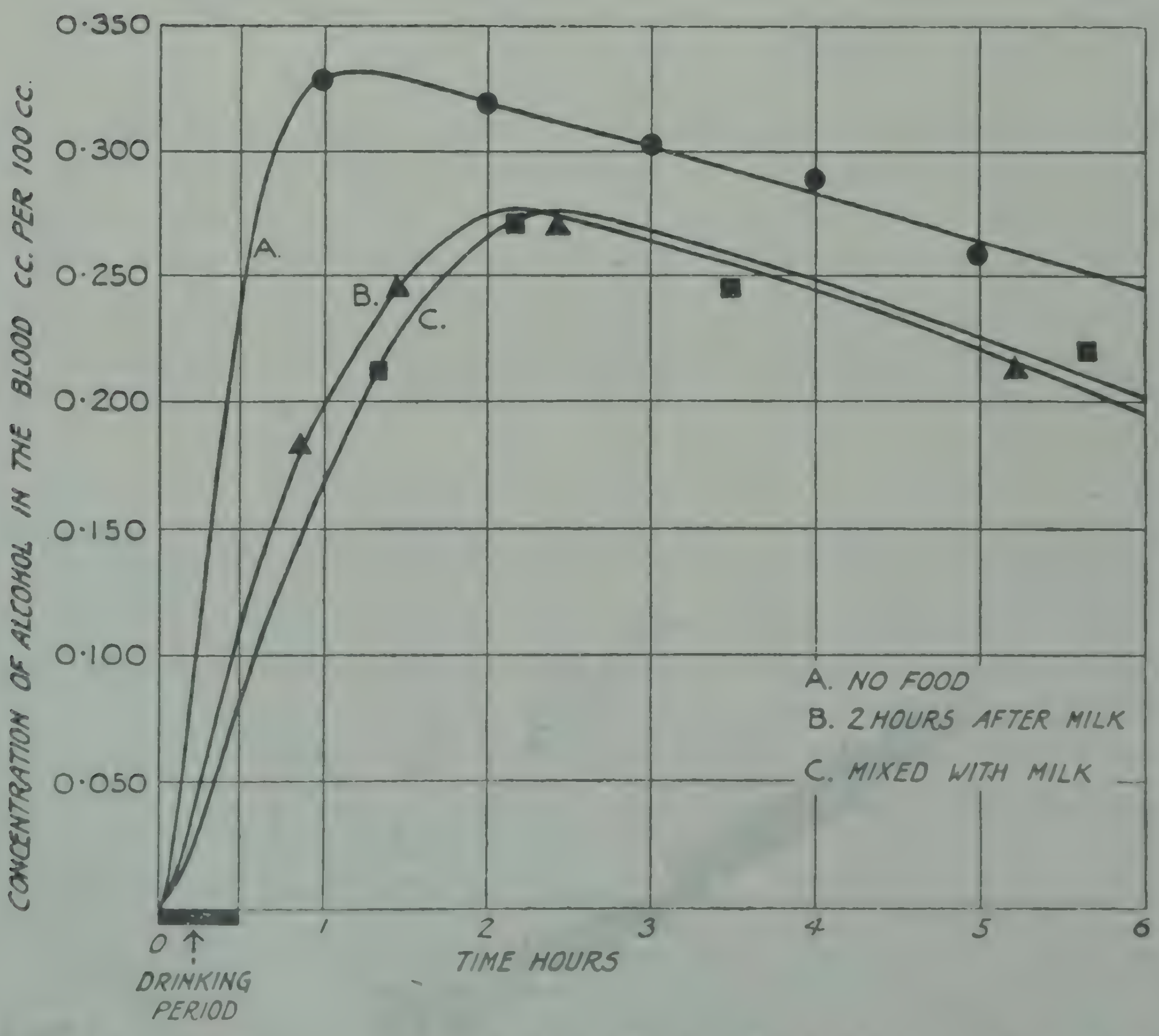


FIG. 542. A. Concentration of alcohol in the blood of a man after taking 171 cc. alcohol in the form of whisky. B. The same dose preceded by 1 pint of milk. C. "The same again," accompanied by a pint of milk. (After MELLANBY.)

effect of any given dose depends on its rate of absorption, which is dependent on the rate of emptying of the stomach, for, although alcohol is absorbed slowly there, absorption is much more rapid in the small intestine. Since the natural rate of absorption, and also the natural tolerance to its action, varies so widely from one subject to another (Fig. 541) no definite degree of intoxication can be attributed to any particular concentration in the blood. The more rapid the increase the greater its effect on the nervous system,* and with rapid absorption some degree of intoxication can be observed at a concentration well below 0·1 per cent., though it is generally stated that 0·15 per cent. usually causes mild intoxication and 0·25 per cent. definite drunkenness. There is no clear evidence of increased tolerance by habituation in man, but such an effect has been shown in dogs.† The taking of food, and particularly of fat, delays the absorption of alcohol (Fig. 542), so that alcohol taken on an empty stomach will produce the speediest and greatest effects.

The rate at which alcohol is removed from the body is more closely related to the size of the liver than to body weight or surface and is affected to some extent by the concentration in the blood ‡ (Fig. 543). It varies also with the general metabolic rate, for it disappears more rapidly after the taking of food, particularly of protein. There is no good evidence that its rate of oxidation is increased by habituation.

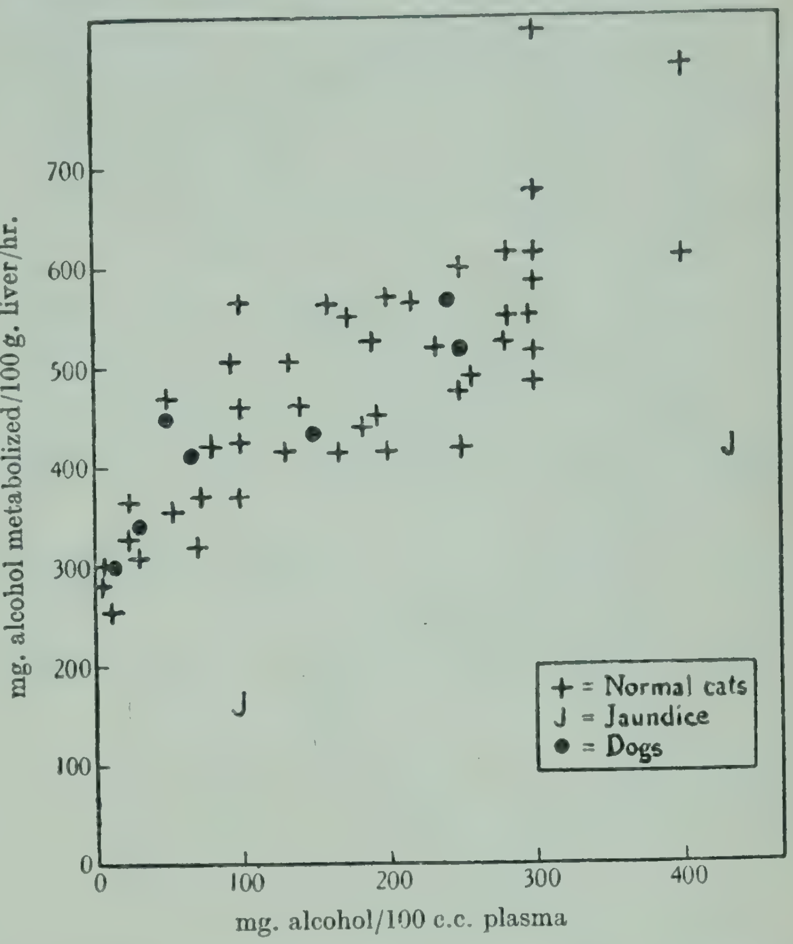


FIG. 543. Relationship between metabolic rate of alcohol per unit liver weight and plasma alcohol concentration. (EGGLETON. *J. Physiol.*, 1940, **98**, 239.)

The removal of alcohol from the blood by excretion and oxidation takes place, according to Mellanby, at a constant rate of about 0·01—0·02 p.c. per hour, so that it takes ten hours to get entirely free from alcohol after it has reached a concentration of 0·12 per cent. in the blood. Alcohol is excreted in the breath and in the urine, the concentration of alcohol in both depending on its concentration in the blood. Alcohol causes a diuresis while the blood alcohol is mounting, and hence a greater one when absorption is slow, because the rise lasts longer : the alcohol in the urine is in equilibrium with that of the plasma water,§ but because the plasma water is lowered by the protein, the concentration in the urine is 1·35 × that in the blood. It is clear from what has been said that the concentration of alcohol in the bladder urine collecting over a long period will bear no close relation to its concentration in the blood at the time of voiding, but a concentration over 0·35 per cent. is likely to correspond to a definite state of drunkenness.

---

* EGGLETON. *Brit. J. Psychol.* (Gen. Sect.), 1941, **32**, 52.
† NEWMAN and LEHMAN. *J. Pharm. Exp. Ther.*, 1938, **62**, 301.
‡ EGGLETON. *J. Physiol.*, 1940, **98**, 239.
§ EGGLETON. *J. Physiol.*, 1942, **101**, 172.

CHAPTER XLIV

THE PHYSIOLOGY OF DIGESTION

CHANGES UNDERGONE BY THE FOODSTUFFS IN THE ALIMENTARY CANAL

In most cases the foodstuffs as we ingest them are colloidal, neutral, and insoluble and have to be rendered soluble and diffusible, in order to allow of their easy transference into the blood. It is necessary that all the carbohydrates shall be altered to the form of monosaccharides. As regards proteins, the change must be even more profound. To form serum albumin, for instance, out of wheat gliadin, an entire reconstruction is necessary. This can only be accomplished by taking the protein molecule to bits, selecting certain of its constituent parts and building these up in the proper proportions to make a new protein molecule. This process of resynthesis must be carried out in the tissue cells themselves.

In all the higher animals the solubilisation of the food is accomplished in the lumen of the alimentary canal by means of *digestive juices,* formed by the activity of glands produced as outgrowths from the wall of the canal. The digestive juices attack the foodstuffs by means of enzymes, and the action of these is hydrolytic. Since each class of foodstuff requires different enzymes, a great variety of these is concerned in the process of digestion.

As the end result of digestion the many kinds of food are reduced to a fairly small number of simple substances. These end-products are :

(1) *Carbohydrates.* The monosaccharides : glucose, fructose, and galactose.
(2) *Fats.* Fatty acids and glycerol ; also finely emulsified fat.
(3) *Proteins.* A great variety of mono- and diamino-acids.

Those constituents of the food, such as the water and salts, which undergo no metabolic alteration in the body, pass along the alimentary canal practically unchanged ; if absorbed they pass without further alteration into the blood.

1. DIGESTION IN THE MOUTH

It is common experience that, when food is taken, there is a flow of " saliva " into the mouth. Saliva is the product of secretion from three pairs of large salivary glands situated in the neighbourhood of the mouth and pouring their secretions into it by means of ducts. It is possible to collect the fluid secreted by each of these glands separately ; and it is found that the saliva varies in properties according to the gland from which it is derived. In addition to these large glands the whole mucous membrane of the mouth is beset with small glands, the buccal glands.

The saliva is in most cases a mixture of the secretions of all these glands. It forms a colourless cloudy liquid, slimy in character. The cloudiness is due to the presence of a number of desquamated epithelial cells, lymphocytes, and clumps of mucin. Its reaction as secreted is neutral, or very nearly so (pH 6·4–7·0), and it is well buffered. On standing or heating it loses CO_2 and becomes alkaline and turbid from deposition of calcium salts, which may be deposited as " tartar " on the teeth.* Its specific gravity varies between 1002 and 1008 and freezing point—0·07 to —0·34° C. Its chief constituents are mucin, and in some cases a diastatic enzyme, ptyalin, and traces of coagulable proteins, potassium thiocyanate, urea, etc. Its approximate composition in man is as follows :

---

* Phillips. *Ztschr. f. physiol. Chem.,* 1935, **233,** 208.

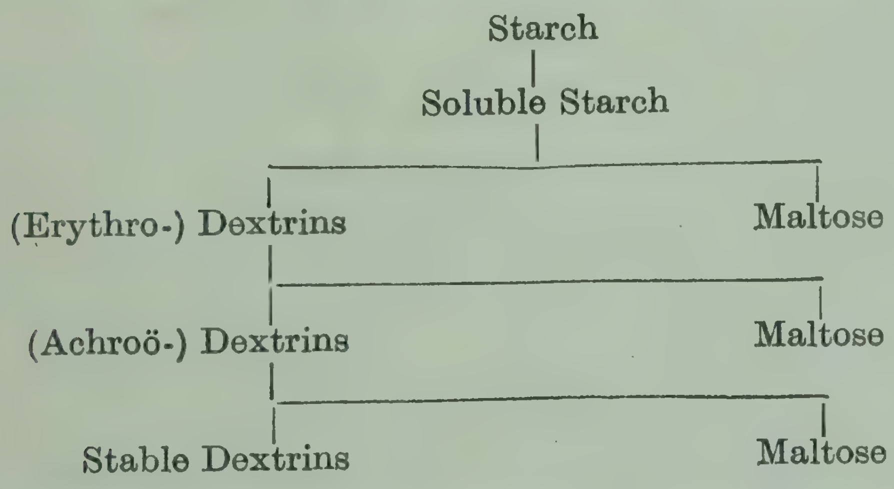
Water	99·27	Na	0·02
Solids	0·73	K	0·06
Organic Solids	0·50	Ca	0·01
Mucin	0·30	CNS	0·03
Inorganic Solids	0·23	HCO_3	0·10
Cl	0·05	Total N.	0·07
P	0·02		

Gases are about 1 c.c. O_2, 2·5 c.c. N_2 and 50 c.c. CO_2 per 100 c.c.

Thiocyanate is formed in the body whenever cyanides or organic nitriles make their appearance, either as the result of administration or as by-products in the normal processes of metabolism. The conversion of the poisonous cyanides into the almost innocuous thiocyanates seems to be a means by which the organism protects itself against the poisonous effects of the former. Saliva of smokers contains more thiocyanate than that of non-smokers, because traces of HCN are absorbed from the smoke. The channels of excretion of the thiocyanate are by the saliva, the urine, and the gastric juice.

The Uses of Saliva. The main function of saliva is to moisten the food and so facilitate its mastication and deglutition. The presence of the mucin is of special value for the latter process, since it renders the mass of food slippery. In animals such as dogs, where the saliva is devoid of any digestive enzyme, this must represent its sole function.

In man and some of the herbivora the saliva contains an enzyme, called *salivary amylase*, or *ptyalin*, which exerts a digestive effect on starch. The stages in the action of saliva on boiled starch can be followed when a very small amount of saliva is added to some starch solution at 37°, and portions of the digest are tested at intervals. The first change is a conversion of the opalescent gelatinising starch solution into a clear solution which no longer sets on cooling, but still gives a blue colour with iodine. This is known as soluble starch. The soluble starch then undergoes hydrolytic dissociation into a dextrin, which gives a red colour with the iodine and is therefore known as erythrodextrin, together with maltose. The erythrodextrin is then hydrolysed into an achroödextrin (giving no colour with iodine) and maltose, and the achroödextrin is still further broken up into stable dextrin and maltose. The conversion of starch into maltose by ptyalin is practically finished when 80 per cent. of the starch has been so converted, the remaining 20 per cent., called stable dextrin, being resistant.* The stages in the conversion are represented in the following Table :



The process by which the huge starch molecule is converted into dextrins and maltose is a very complicated one, and a number of intermediate compounds of dextrins and maltose may exist between those whose presence is revealed by their varying reaction to iodine. On very prolonged action of saliva, some of the maltose formed is further hydrolysed to glucose, but this action is too slow to be of any importance in the body.

Ptyalin is most active at a *p*H of about 6·0, so that the addition of minute traces of acid to an alkaline saliva increases its activity. Further acidification, however, rapidly destroys it. Removal of salts by dialysis renders it almost or quite inactive, the activity returning when the salts are restored. Addition of small amounts of chlorides greatly accelerates the action of saliva. It acts most rapidly at 45°, but even at 0° C. its action is still just perceptible. If heated to 60° C. the amylase is destroyed.

---

* Lovatt Evans. *Journ. Physiol.*, 1912, **44**, 191.

When the starch is in the solid condition, as in biscuits and most farinaceous foods, its stay in the mouth during the normal process of mastication is not long enough to allow any considerable hydrolysis to occur. When a meal is taken, the food is swallowed and forms a mass which is penetrated only slowly by the acid gastric juice secreted by the stomach within five minutes of the taking of food. Even half an hour after a meal the interior of the mass of food in the stomach may be still found to be neutral or slightly alkaline. The food, therefore, thoroughly mixed with saliva, remains in the stomach for thirty to forty minutes, gradually reaching its optimum *p*H—and with chloride accumulating and also accelerating its action—before the amylase is destroyed by the acid gastric juice. During this time the amylase continues to act, so that the chief part of the salivary digestion occurs actually in the stomach, and results in extensive conversion of the starch into dextrins and maltose.*

The Secretion of Saliva. The greater part of the saliva is formed in man by the sublingual and the submaxillary (sub-mandibular) glands.

The arrangement of these glands, especially of those in the floor of the mouth, varies somewhat in different animals. They can be divided into three classes. Among the small buccal glands we may distinguish two types, the *mucous gland* and the *serous gland* (Fig. 544). The mucous gland is distinguished by the fact that its short duct opens into wide alveoli, the lining cells of which are distended with mucin and therefore present a clear unstained space in the section. In the serous gland, the duct lined with columnar cells branches into a series of acini, which present a well-marked lumen and are lined with small granular cells with a very distinct and well-staining nucleus. The same general distinction can be made out in the large salivary glands. The parotid gland in man and in all the higher mammalia is a typical serous gland, though here and there a mucous cell may be seen. The sublingual and submaxillary glands in man represent a third type called *mixed* glands. Most of the alveoli are mucous in character. At the ends of the alveoli are seen crescent-shaped serous cells between the mucin-distended cells and the basement membrane. These are known as the *demilune cells*, or the *Crescent cells* of Gianuzzi. In some cases these mucous alveoli with demilunes may be found alongside of typical serous alveoli.

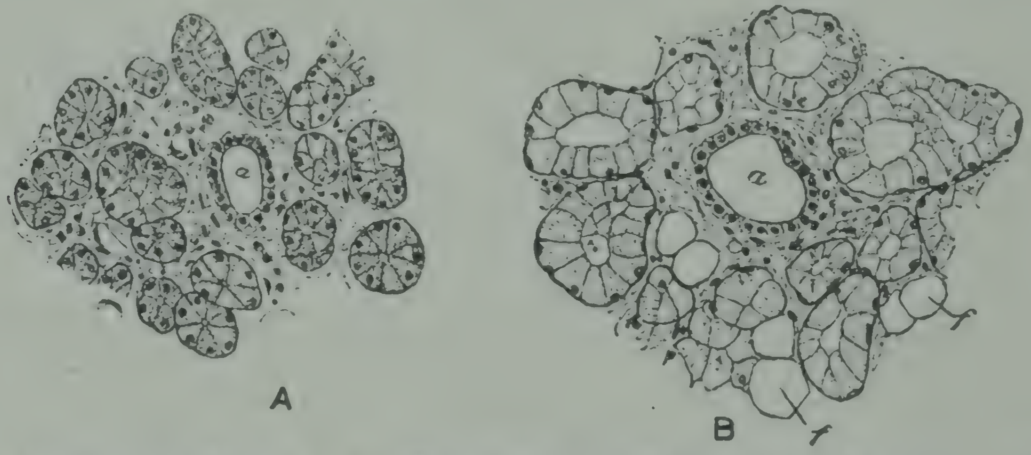


FIG. 544. A. Serous Gland. B. Pure Mucous Gland from Mouth. (KÖLLIKER.)
a. Ducts. *f*. Fat cells.

The secretion of the separate glands in man can be studied by passing a fine catheter into the appropriate duct, and also in certain rare cases where fistulæ occur in the parotid duct. The normal behaviour of the salivary glands is best studied by the aid of fistulæ. It is possible to transplant the papilla, on which the duct opens, to the outside, so that the saliva from any particular gland shall flow externally instead of into the cavity of the mouth. By attaching a small funnel to the fistulous opening, it is easy to collect the pure saliva. These fistulated glands retain their normal structure and functions for many months, though the gland finally becomes fibrotic.

In man the glands are continually secreting slowly † ; in carnivora there is normally no secretion when no stimulus such as food is present ; as soon as food is introduced into the mouth there is a rapid secretion of saliva, the relative extent to which different glands are involved varying according to the nature of the stimulation. Thus with meat and other moist foods there is (in the dog) only a small amount of secretion, which is derived chiefly from the submaxillary and sublingual glands, and is rich in organic constituents. When dry or undesirable material, such as dry powdered meat, or sand, is introduced the flow of juice is more copious and more watery. Secretion may be produced reflexly in the dog by showing food, or even by the suggestion that dry powder is to be introduced into the mouth. This is an example of a conditioned reflex.

---

* BEAZELL. *Amer. J. Physiol.*, 1941, **132**, 42.

† LASHLEY. *J. exp. Psychol.*, 1916, **1**, 461.

The serous and mucous glands differ in the nature of their secretion. A serous gland, such as the parotid, gives a clear watery secretion almost free from mucin, but containing small traces of coagulable protein. The mucous gland delivers a secretion which is viscid from the presence of mucin, and contains also a small trace of coagulable protein. In man amylase is found in the saliva from both kinds of glands, though it is about four times as concentrated in that secreted by the parotid gland. The total amount of saliva which may be obtained varies of course in different animals. Each gland may, however, in the course of the day give an amount of juice far exceeding, *e.g.* ten or twelve times, its own weight. In man it is reckoned that over one litre of saliva may be formed every twenty-four hours, and in the herbivora such as the horse, the total diurnal production must amount to many litres. In thirsting the reflex secretion of saliva is reduced.

NERVOUS ARRANGEMENTS OF SECRETION. Activity of the salivary glands must be elicitable by reflex means. The *afferent nerves* in this reflex include those that supply the mucous membrane of the mouth, *i.e.* the fifth nerve, the pharyngeal branch of the vagus, and the glossopharyngeal.

The salivary " *centre* " consists of a neurone pool in the reticular formation extending from the sensory nucleus of the VIIth cranial nerve down to the anterior end of the nucleus ambiguus ; the more caudal part supplies the parotid, the rest the other two glands. Certain areas of the cerebral cortex also give salivation when excited, but the significance of this is doubtful.

Sympathetic subordinate centres are located in the spinal cord, and are probably under superior control from the hypothalamus.

Each one of the large salivary glands receives efferent nerve fibres from two sources : from the parasympathetic and from the sympathetic system.

The *parasympathetic fibres* to the submaxillary, the sublingual, and the retrolingual glands pass into the facial nerve, and from this nerve along the chorda tympani to the lingual division of the fifth nerve.* The lingual nerve passes below the duct, and just before it crosses the two ducts of the submaxillary and retro- or sub-lingual glands, it gives off a branch of *chorda tympani fibres*, which runs backwards along the submaxillary duct to be distributed to the glands, and in its course gives off fibres also to the retrolingual (Fig. 545). Two neurones are involved in the cranial supply. By means of the nicotine method, Langley showed † that all the preganglionic fibres to the sublingual and the submaxillary glands end somewhere near the glands in connection with ganglion cells, from which non-medullated post-ganglionic fibres pass to the glands. These relay ganglia for the submaxillary gland are scattered cells lying in " Langley's ganglion," in the hilum of the gland itself ; in the cat those for the retrolingual gland are in the so-called " submaxillary ganglion." The fibres to the sublingual gland in man probably take a similar course. The fibres are apparently finally distributed to the secreting alveoli, where they end in the mucous secreting cells.‡

The parasympathetic fibres to the parotid gland leave the medulla by the glossopharyngeal nerve and run in its tympanic branch (nerve of Jacobson) to the tympanic plexus, and then by the small superficial petrosal nerve to the otic ganglion, where relays occur, the post-ganglionic fibres then passing to the gland by the auriculo-temporal branch of the third division of the fifth nerve.* The site of termination in the gland is doubtful.

---

* LUDWIG. *Ztschr. f. rat. Med.*, 1851, N.F. 255 ; CLAUDE BERNARD, *Syst. Nerveux*, 1858, II., 150.

† LANGLEY. *Journ. Physiol.*, 1890, **11**, 123.

‡ RAWLINSON. *Anat. Rec.*, 1933, **57**, 289.

The *sympathetic supply* to all these glands is contained in fine filaments on the walls of the arteries with which they are supplied. The fibres are derived from the spinal cord, whence they issue by the upper three or four ventral thoracic nerve roots. They pass into the cervical sympathetic as preganglionic fibres which end around the cells of the superior cervical ganglion, and fresh

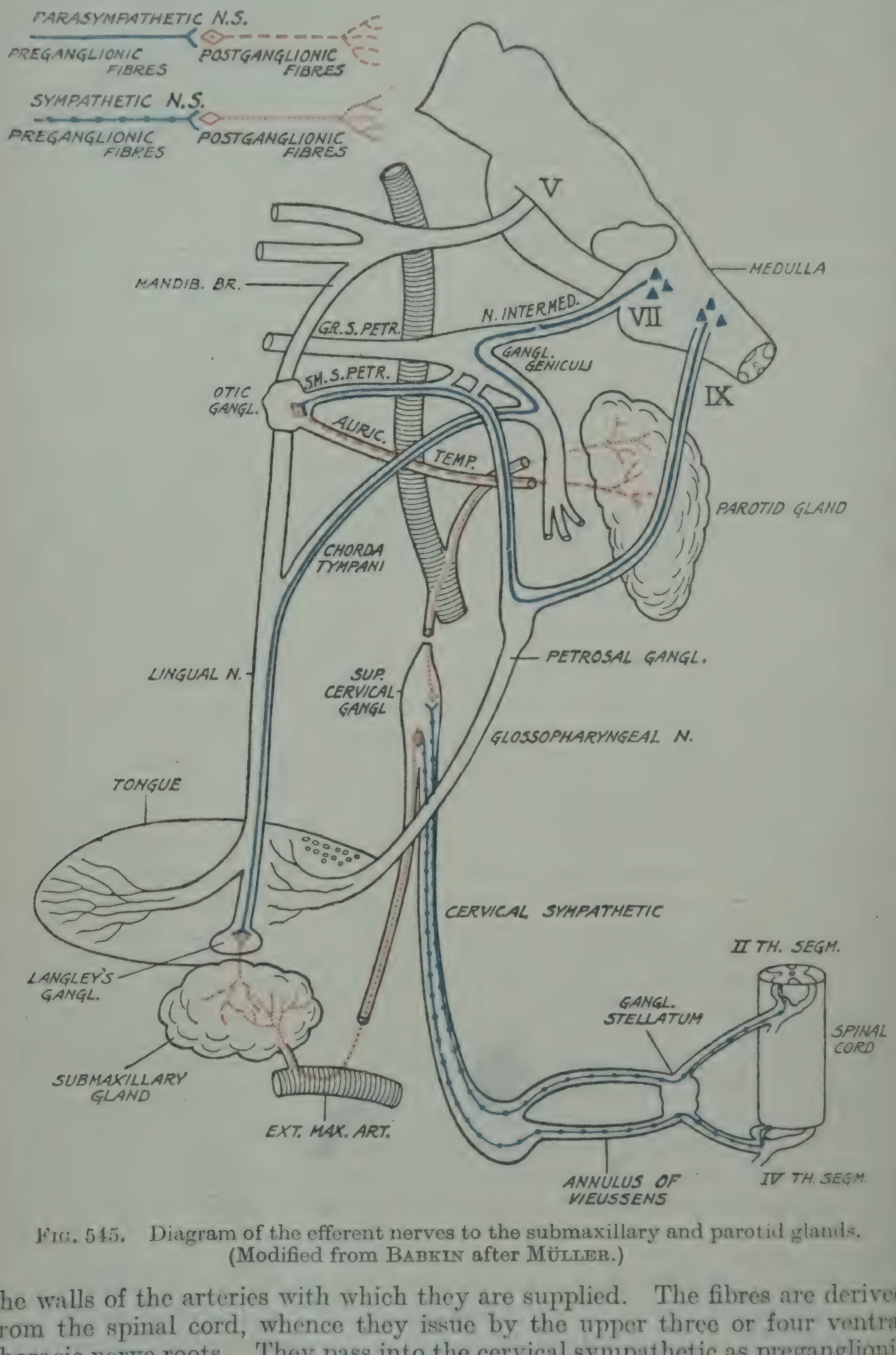

PARASYMPATHETIC N.S.
PREGANGLIONIC FIBRES
POSTGANGLIONIC FIBRES
SYMPATHETIC N.S.
PREGANGLIONIC FIBRES
POSTGANGLIONIC FIBRES
V
MEDULLA
MANDIB. BR.
N. INTERMED.
VII
GR. S. PETR.
GANGL. GENICULI
SM. S. PETR.
OTIC GANGL.
IX
AURIC.
TEMP.
PAROTID GLAND
CHORDA TYMPANI
PETROSAL GANGL.
LINGUAL N.
SUP. CERVICAL GANGL
GLOSSOPHARYNGEAL N.
TONGUE
CERVICAL SYMPATHETIC
II TH. SEGM.
LANGLEY'S GANGL.
GANGL. STELLATUM
SPINAL CORD
SUBMAXILLARY GLAND
EXT. MAX. ART.
ANNULUS OF VIEUSSENS
IV TH. SEGM.

Fig. 545. Diagram of the efferent nerves to the submaxillary and parotid glands. (Modified from Babkin after Müller.)

---

* Claude Bernard. *Leçons sur la physiol. et la pathol. dy système nerveux*, Paris, 1858, II., p. 158.

relays of fibres, chiefly non-medullated, arise from these cells and travel on the walls of the branches of the external carotid artery to their destination, which is in the demilune cells of mixed glands, and the main cells of serous glands.

The effects of stimulating the peripheral ends of the parasympathetic nerves going to the glands present a general resemblance, whichever be the gland involved.* Within a period of half to two seconds after the stimulation has been applied, a secretion of saliva is produced, presenting similar characters to that which would be obtained from the gland under normal conditions. The concentration of the saliva, as well as its rate of secretion, depends on the strength of the stimulus. The following table (Heidenhain) shows the effect on the amount and composition of submaxillary saliva obtained by weak and strong stimulation of the chorda tympani nerve :

Strength of stimulus.	Quantity in one minute.	Per cent. of organic solids.	Per cent. of salts.
Weak . .	0·17	0·84	0·20
Strong . .	0·72	2·06	0·46
Weak . .	0·17	1·67	0·26

With the strong stimulus the amount of saliva was increased over fourfold, while the percentage of solids in the saliva was raised. If the excitation be continued for a considerable time, there is a gradual rise in the percentage of inorganic salts and a fall in the percentage of organic matter.

The cranial nerves going to these glands produce another important effect, namely vaso-dilatation. It was shown by Claude Bernard that on exciting the chorda tympani the flow from the vein of the submaxillary gland might increase four to eight times, and indeed to such an extent that, in spite of the fact that the secreting gland has a greatly increased oxygen utilisation,† the blood passing through the gland did not stay long enough to lose much oxygen. As the result of exciting the chorda, therefore, the blood coming from the vein may show distinct pulsation, and may have a brilliant scarlet hue just as if it were derived from an artery. The same dilatation has been observed to attend excitation of the cranial supply to the parotid gland.

It is probable that all the effects of stimulation of the parasympathetic nerve supply upon a salivary gland are due to the liberation at the nerve terminations‡ of small amounts of acetyl choline, the injection of which into the blood stream produces effects similar to those of the nerve stimulation. Both effects are prolonged by eserine and abolished by atropine. Pilocarpine causes profuse salivation, but the saliva produced differs in composition from that resulting from chorda stimulation.

The effects of exciting the sympathetic nerve supply differ according to the gland and the animal. Always there is vaso-constriction. In the dog, excitation of the cervical sympathetic causes the secretion of a viscid saliva from the submaxillary gland but no secretion is obtained at all from the parotid gland, though there are histological changes in the gland cells. In the cat, the submaxillary saliva obtained on sympathetic excitation may be as copious as that obtained on stimulation of the chorda tympani. We shall have later on to discuss how far these results are to be ascribed to a fundamental difference in the effects of impulses carried by the two sets of nerve fibres, and how far to their different effects on the blood vessels.

A motor effect, producing a shrinkage of the gland, is also seen in all the salivary glands on stimulation of the sympathetic supply. It is caused by the contraction of special myo-epithelial cells lying just inside the basement membrane.§

---

* HEIDENHAIN. *Pflüger's Arch.*, 1878, **17**, 1.

† J. BARCROFT. *Journ. Physiol.*, 1901, **27**, 31.

‡ BABKIN, GIBBS and WOLFF, *Arch. f. exp. Path.*, 1932, **168**, 32 ; GIBBS and SZELÖCZEY, *ibid.*, p. 64.

§ BABKIN and MACKAY. *Amer. J. Physiol.*, 1930, **91**, 370.

CHANGES IN THE GLAND ACCOMPANYING SECRETION. The fact that a submaxillary gland of the dog under favourable conditions will secrete its own weight of saliva in five minutes, and will on continued stimulation secrete for many hours, shows that there is a continual renewal of the fluid which is turned out in the secretion. If we refer to the diagrammatic representation of the elements which make up a secreting lobule and which may be involved in the act of secretion (Figs. 546, 547), we see that between the lumen of the duct and the blood, which must be regarded as the source of the fluid, the following layers of cells intervene: (1) the endothelium of the blood capillaries; (2) the basement membrane; (3) the epithelial cells of the gland proper. We have in the first place to decide to which of these elements can be ascribed the chief part in the act of secretion.

The secretory activity is usually associated with a considerable dilatation of the vessels of the gland, and a consequent large increase (three and eight times) in the blood flow through the gland. Such an increase in the supply of blood affords a source for the large quantity of fluid which is turned out in the saliva, and raises the pressure in the capillaries of the gland. We cannot, however, regard this rise of pressure as the essential factor in the act of secretion. If a small dose of atropine be administered, excitation of the chorda causes the same vaso-dilatation as before, though no secretion is produced. Moreover, Ludwig showed that if two manometers be connected, one with the carotid artery and the other with the duct, on stimulating the chorda tympani nerve the mercury will be driven up by the pressure of the secretion in the corresponding manometer until it attains a height which may be double that in the carotid artery,* and therefore must be still greater than the pressure in the capillaries of the gland. This showed the impossibility of the act of secretion being a process of mere filtration.

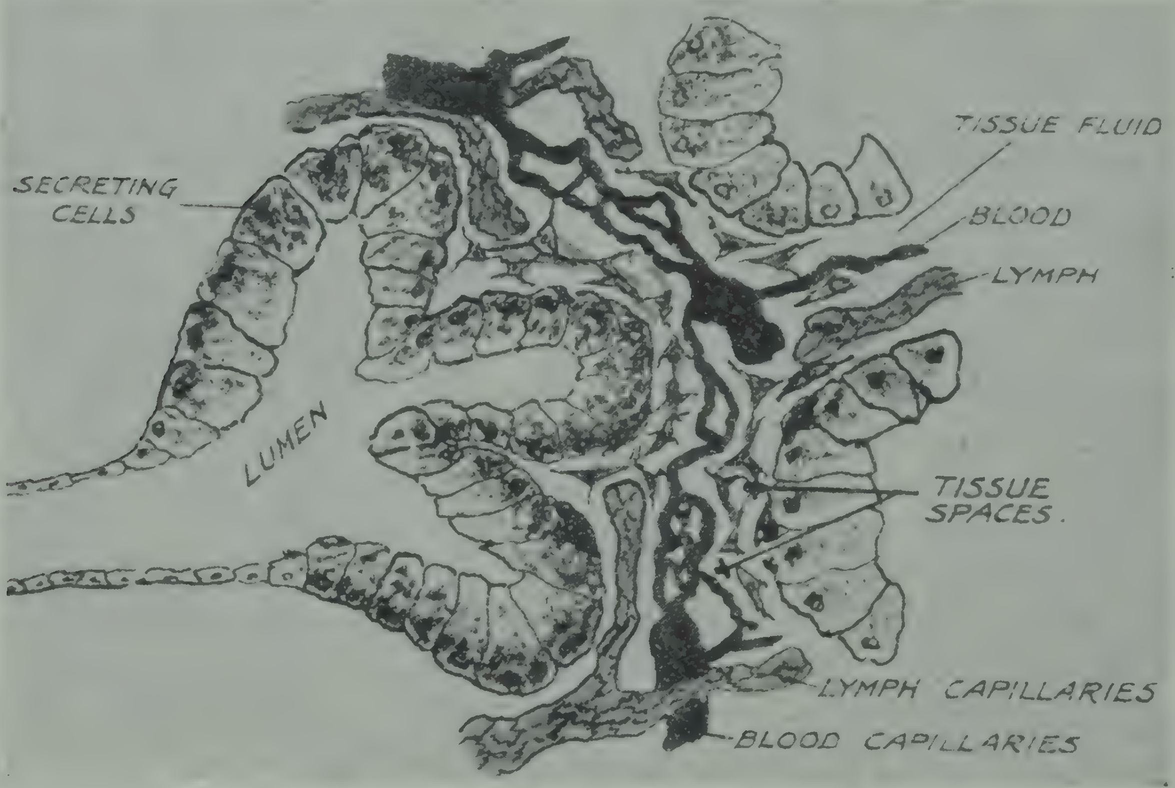


FIG. 546.

There is further evidence that work is done in the production of the salivary secretion. When a fluid containing salts in solution is filtered through a porous membrane, the filtrate has the same content in salts as the original fluid. We can effect a separation of dissolved salts from solvent by filtration under pressure through a semipermeable membrane. Under these circumstances a very large pressure is necessary in order to cause the filtration of any fluid at all, a pressure greater than the osmotic pressure exerted by the substances in solution. Thus if we were filtering a 1 per cent. solution of NaCl through a semipermeable membrane, we should have to exert a pressure of about seven atmospheres in order to obtain a filtrate free from sodium chloride.

---

* LUDWIG. *Mitt. d. Zürch. Naturf.*, 1851, Nr. 50, 271. (Ostwald's *Klassiker d. Exakten, Wiss.* Nr. 18.) This effect is only well seen if the arterial pressure is low.

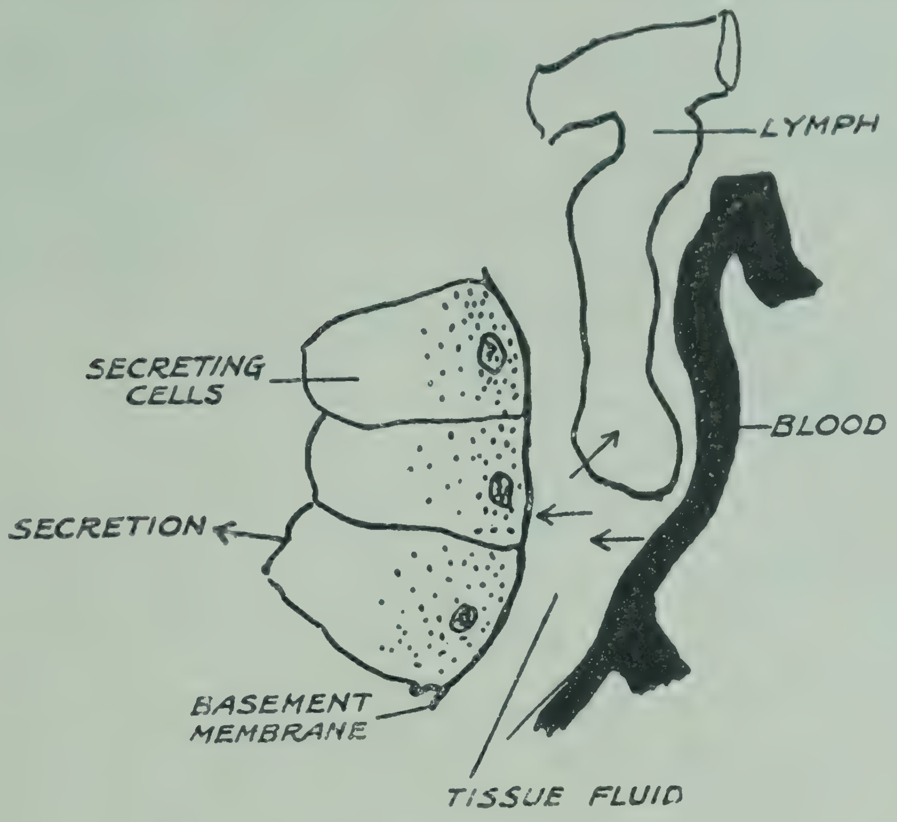


FIG. 547.

FIGS. 546 and 547. Diagrams to show Relation of the Secreting Cells of a Gland to the Blood and Tissue fluid and Lymph Drainage. Figure 547 is schematic.

To obtain a filtrate containing half the amount of sodium chloride, if such were possible, would therefore need a pressure of about three and a half atmospheres. On comparing the osmotic pressures of saliva and blood respectively—using the depression of freezing-point as our index—we find that the osmotic pressure of saliva is between half and three-quarters that of the blood plasma. Supposing the membrane separating the lumen of the duct from the blood vessels could be regarded as a semi-permeable membrane, we should need, in order to effect the separation, a pressure ten to twenty times as large as the arterial blood pressure. We must conclude, then, that work, both osmotic and mechanical, is performed in the separation of the fluid from the blood and its transference in the form of saliva to the duct.

A simple experiment shows that this work must be effected by the gland cells themselves. The fluid passes from the blood vessels first into the tissue spaces, whence it is taken up by the secretory cells. If the first act in secretion consisted in an increased exudation of fluid from the blood into the tissue spaces, the first effect of exciting the chorda tympani nerve should be to cause the swelling of the whole gland. If all secretion be prevented by the previous administration of atropine, stimulation of this nerve produces, as might be anticipated, an increased volume of the gland in consequence of the dilatation of its vessels. If, however, the gland be allowed to secrete, we obtain, in spite of the simultaneous increase in size of the vessels, an actual diminution in the size of the

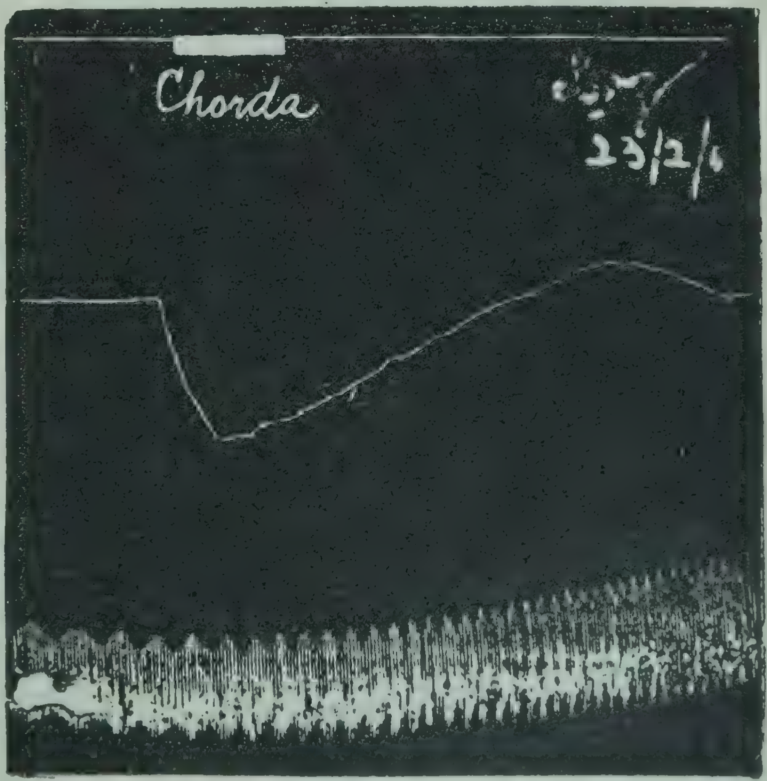


FIG. 548. Tracing of Volume of Submaxillary Gland, showing Decrease on Excitation of Chorda. (BUNCH.)

gland itself,* showing that the first effect of the stimulation is on the cells of the alveoli (Fig. 548). Under the stimulus, these empty themselves of some of their contents, replenishing themselves later at the expense of the tissue fluid. The increased passage of fluid from blood to tissue space is therefore a secondary and not a primary effect of the nerve stimulation.

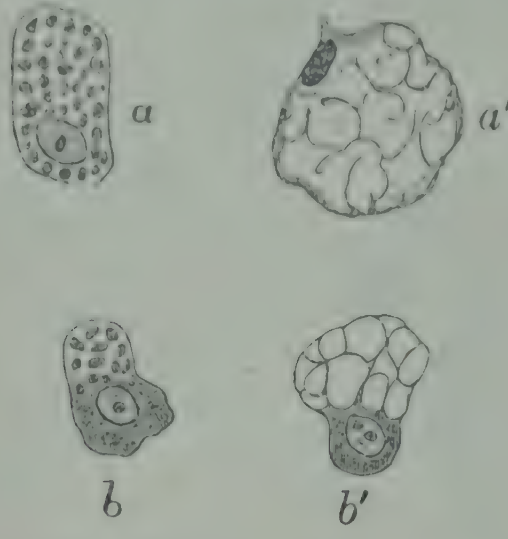


Fig. 549. Mucous Cells from a Fresh Submaxillary Gland of a Dog. (Langley.)

a. Mucous cell examined fresh from a resting gland.
a′. The same cell treated with weak alcohol.
b and *b′*. Cells from a discharged gland before and after treatment with weak alcohol.

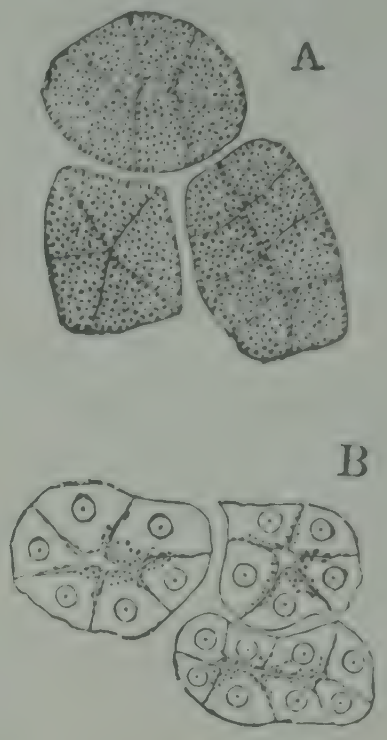


Fig. 550. Acini of a Serous Salivary Gland. (Langley.)

A. Resting condition.
B. Discharged condition.

Histological Changes during Secretion. The process of secretion is associated with changes in the structure of the secretory cells. The changes are of the same general character whatever class of glands we investigate, though the ease with which they are to be demonstrated varies with the effects of the fixing fluids employed. If a small fragment of a mucous gland be teased in blood serum or in 2 per cent. NaCl solution, the cells are found to be packed with a mass of coarse, highly refractive granules (Fig. 549). If a corresponding specimen be made from a serous gland (Fig. 550), the cells are also packed with granules which, however, are much finer in structure. On making similar specimens from glands which have been forced to secrete for six or seven hours, the individual cells are found to be much smaller and the protoplasm of the cell is relatively increased in amount, while the granules are much fewer and are now confined almost entirely to the inner margin of the cell.† Activity is thus associated certainly with a discharge of granules, and probably with some increased building up of protoplasm. We may regard the act of secretion as determined by the alteration of the granules and their discharge, together with water and salts, to form the specific secretion of the gland. During rest the granules are re-formed by precipitation in, or modification of, the protoplasm surrounding the nucleus. We have evidence that, although the granules enter the secretion, they represent, not the secretion itself, but a precursor of some

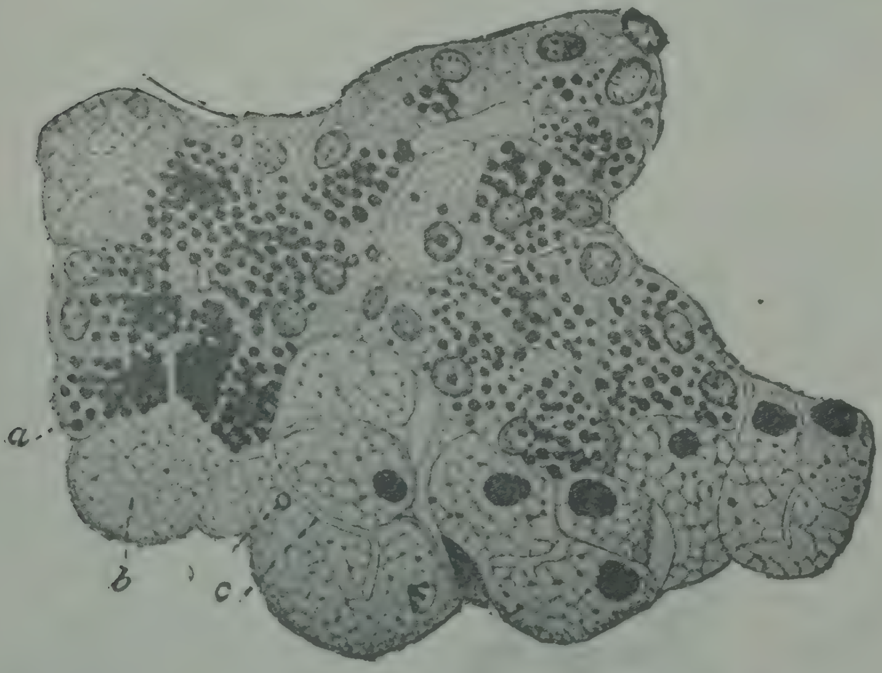


Fig. 551. Submaxillary Gland of Rabbit. (Schafer after E. Müller.)

The cells, all serous, are in different functional states:
a. A loaded cell.
b. A discharged cell.
c. A secretory canaliculus penetrating into a cell.

---

* Bunch. *J. Physiol.*, 1900, **26**, 1.
† Langley. *Journ. Physiol.*, 1879, **2**, 261.

at any rate of its constituents. Thus if acetic acid be added to the saliva obtained from the submaxillary gland, the mucin is precipitated as threads and films. If the granules in the secreting cells also consist of mucin, we should expect acetic acid to have a coagulating effect upon them. We find on the contrary that, on allowing acetic acid to flow over a section of the fresh gland, the granules at once swell up and burst. We must regard these granules therefore, not as mucin, but as a precursor of mucin, mucinogen. The effect of ordinary fixing reagents is to cause these granules to swell up so that the cells become filled with a mass of mucin giving the typical hyaline appearance of ordinary sections of these glands. In the case of the serous glands the granules (Fig. 551) are apparently protein in nature. Where ptyalin is a constituent of the saliva, we are probably justified in assuming that it is contained as a precursor in the granules. It is customary to speak of the granules in a secreting gland as *zymogen* granules, *i.e.* the precursors of enzymes. It is probable that we ought to regard these granules, not merely as precursors of the constituents of the secretion, but as cell organs in which proceed a whole series of chemical and osmotic changes, which determine the production of the fully formed secretion directly from the protoplasm and indirectly from the ordinary constituents of the surrounding tissue fluid.

ELECTRICAL CHANGES. The interpretation of the electric potential changes which accompany secretion is difficult, owing to the simultaneous effect of the movement of fluids through porous membranes. If the hilum of the submaxillary gland and its outer surface be connected with a galvanometer, the outer surface is positive to the hilum. On exciting the chorda tympani nerve a diphasic effect is generally obtained, the resting difference being first increased and later on diminished. On excitation of the sympathetic nerve we generally obtain a purely negative variation of the resting difference. These results were interpreted by Bayliss and Bradford * as due to the co-operation of the two factors, chemical change in the gland cells and movement of fluid through the cells. The positive variation, *i.e.* the current from within out, was ascribed to the movement of fluid, whereas the negative variation of the resting difference was thought to be an action potential, due to the chemical changes in the gland cells.

THE SIGNIFICANCE OF THE DOUBLE NERVE SUPPLY. Heidenhain † claimed that stimulation of the sympathetic supply to a salivary gland causes histological changes in the gland even more marked than those produced by the cranial nerve. He therefore put forward the hypothesis that the salivary glands are supplied by two fundamentally different classes of fibres, namely: (1) trophic fibres, which are responsible for the production of the specific constituents of the secretion, and (2) secretory fibres, excitation of which causes the cells to take up water and salts from the tissue fluid and blood, and pass them in large quantities into the duct. The sympathetic supply would consist almost entirely of trophic fibres, whereas secretory fibres would predominate in the parasympathetic supply.

According to Langley‡ and Carlson, the difference in the action of the chorda tympani and of the sympathetic on the submaxillary gland is due to the opposite action of these nerves on the blood supply to the gland. But this is inadequate to explain the differences observed in the secretion of saliva reflexly aroused by introduction of substances into the mouth. In a dog a copious flow of saliva may be caused by the introduction of 0·25 per cent. hydrochloric acid or of meat powder. The amount of saliva secreted under the two circumstances is approximately the same, but that evoked by the introduction of meat powder contains about five times as much organic solids as that which follows the introduction of acid into the mouth. The clue to the variation is probably to be sought in the different effects of sympathetic and parasympathetic fibres (or of their chemical representatives adrenaline and acetyl choline) upon different cells of the salivary glands. Stimulation of the chorda in the cat, for instance, causes profound histological changes in the alveolar cells of the submaxillary gland, while sympathetic stimulation, leaving those unaffected, causes changes in the demilune cells.§ The composition of the salivas obtained in the two cases is characteristic of the secretions of the mucous and serous cells respectively.||

Even this explanation may not be complete; Babkin has shown that variation in the saliva is found after complete section of the sympathetic, and that there is the same acceleration of the circulation through the gland whether the secretion is aroused by introduction of meat powder or of acid. It is impossible, therefore, to explain the difference in the composition of the saliva obtained under these two circumstances as due solely to differences in the blood supply to the gland, and we must conclude that the chorda tympani contains different kinds of fibres which are excited to varying extents according to the nature of the reflex stimulation. At any rate, it is certain that, according to the nature of the reflex stimulus, either the secretion of water and salts or the secretion of organic solids may preponderate, altogether apart from changes in the circulation simultaneously evoked. It has also been found that simultaneous stimulation of the chorda tympani and sympathetic nerves gives rise to an *augmented secretion* of saliva, greater than that given by the chorda alone, although the blood flow through the gland is less. (GESELL.)

---

* BAYLISS and BRADFORD. *J. Physiol.*, 1885, **6**, xiii; 1886, **7**, 217.
† HEIDENHAIN. *Pflüger's Arch.*, 1878, **17**, 1.
‡ LANGLEY. *J. Physiol.*, 1878, **1**, 96.
§ RAWLINSON. *Anat. Rec.*, 1933, **57**, 289.
|| LANGSTROTH, MCRAE and STAVRAKY. *Arch. internat. pharmac.*, 1938, **58**, 61.

The Energy Involved in the Act of Secretion. Barcroft [*] attempted to determine the total amount of energy set free by the gland in secretion, by measuring its respiratory exchange under conditions of rest and activity. He found that the resting submaxillary gland in a small dog took up 0·25 c.c. of oxygen per minute, while during active secretion it absorbed 0·86 c.c. O_2. Assuming that the total oxygen taken up is employed in the oxidation of a food substance, such as glucose, and that the whole of the energy of the chemical changes is set free in the form of heat, we find that a resting gland weighing about 6 grammes produces about 1·1 calories per minute. We know, however, that a certain amount of external work is performed in the secretion of a saliva of lower osmotic pressure than the blood and also, when there is any resistance to the flow of saliva through the duct, in raising the hydrostatic pressure of the saliva in the duct.

Even during rest, changes are going on in the gland cells, changes leading to the formation of granules. When activity is excited by the stimulation of secretory nerves, the primary change appears to involve simply the granules. These absorb water, apparently against the osmotic pressure of the surrounding cytoplasm. Those nearest the lumen swell up, become converted into spheres containing water and salts in smaller proportion than exists in the fluid bathing the cells (and presumably in the protoplasm surrounding the granules), and in this swollen form are discharged or ruptured on the periphery of the cell into the lumen, so giving rise to secretion.

This discharge of a fluid with a smaller molecular concentration than the blood plasma must lead to an increased osmotic pressure in the remaining parts of the cell. This would induce a flow of water from tissue fluid into cell, and the consequent concentration of the tissue fluid would in the same way cause a flow of water into it from the blood. This pull of water from the blood by the cell is still further increased in another way. The act of secretion can be carried out only at the expense of chemical changes in the cell, which will result in the formation of a number of small molecules from the great colloid molecules of the protoplasm. The products of metabolism, or *metabolites*, will therefore accumulate in the cell, pass into the tissue fluid, and increase the osmotic pressure of the latter. This will call forth an increased transudation of fluid from the blood vessels, and will therefore produce a distension of the tissue spaces of the gland and an increased discharge of lymph along its efferent lymphatics.

As a secondary result of the activity there is a tendency to increased assimilatory changes and a preparation of the cell for further secretory changes, either immediately or thereafter. During prolonged secretion the mucin content of the saliva steadily diminishes. The non-mucin nitrogen of the saliva is, however, not reduced, and hence is being continually passed on from the blood. Replenishment of mucinogen after complete exhaustion may take as long as three days. The submaxillary gland uses glucose from the blood stream at an increased rate when the gland secretes. (Anrep and Cannan.[†])

In the gland, as in muscle, when we attempt to form a conception of the mechanism of the chemical machine in the living cell, we are brought up against insuperable difficulties, and any theories so far put forward merely serve to show how far we are still from the actual comprehension of the events occurring in every living cell and underlying its conditions of rest and activity.

For Reference

Babkin (1938). The Physiology of the Salivary Glands, Chap. VI, pp. 219-249, of Dental Science and Dental Art. Ed. Samuel M. Gordon. (Philadelphia: Lea and Febiger.)

2. MASTICATION AND DEGLUTITION

Mastication. In the act of mastication the food is mixed with saliva and ground by the teeth, aided by the cheeks, lips and tongue. Salivary secretion is augmented by mastication, and the mass of food, moistened and lubricated by the saliva, is finally rolled between the tongue and hard palate to form a bolus, preparatory to swallowing.

It has been shown by Sherrington [‡] that mastication may consist of a series of reflexes, the closure of the jaws producing a stimulus to reflex opening; then, by rebound of the inhibited jaw-closure, another closure results, and so on, so long as anything biteable remains between the jaws.

Deglutition. The food after mastication is carried to the stomach by a complex series of co-ordinated movements involving the muscles of the mouth, pharynx and œsophagus (Fig. 552).

Various methods have been used to study the process of deglutition in man and the higher animals. In various animals, *e.g.* ostrich, horse, giraffe, the movement of the bolus down the œsophagus can be seen or felt from the outside of the neck. Important information can be obtained by allowing a man or animal to swallow either a fluid or solid with which barium sulphate is mixed, and observing the passage of the opaque substance under the X-rays. The substance may be mixed with milk for a fluid, or with bread and milk for a semi-solid substance, or may be enclosed in a cachet and swallowed as a solid bolus. The various sounds produced in deglutition may be timed and suitably recorded in relation to other events. The relative time relations of the events at different parts of the mouth, pharynx and œsophagus may be obtained by passing tubes either open, or provided with rubber balloons, to different levels, that for the pharynx being passed in through the nasal passage, and connecting these sounds with recording tambours or piston recorders. This method has been employed both in men and in animals, but the records from balloons are not easily interpreted because a stationary balloon would probably be treated like a very large bolus, and be subjected to powerful efforts to move it onwards.

---

[*] Barcroft. *J. Physiol.*, 1901, **27**, 31.

[†] Anrep and Cannan. *J. Physiol.*, 1922, **56**, 248.

[‡] Sherrington. *J. Physiol.*, 1917, **51**, 420.

Observations of the movements of the soft palate and walls of the pharynx have been made also on patients suffering from wounds which exposed those structures to view.*

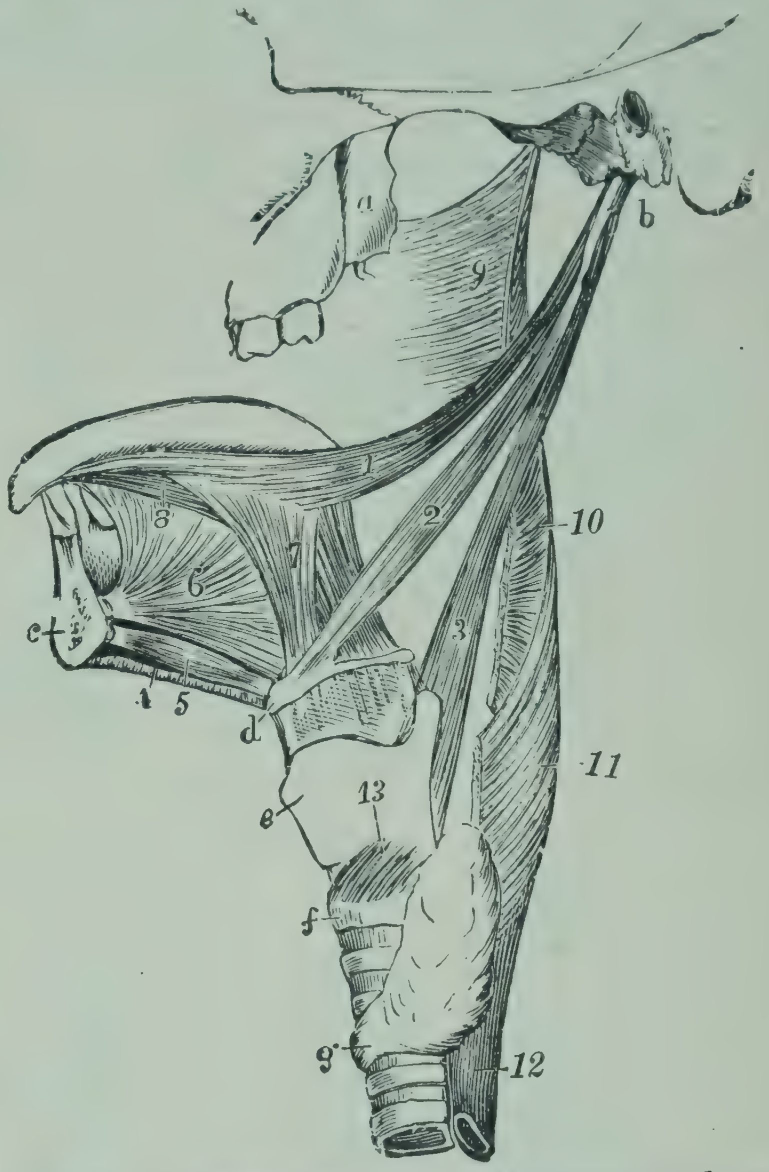


Fig. 552. Dissection to show Muscles employed in Deglutition.

b, styloid process, from which arise 1, the styloglossus; 2, the stylo-hyoid; 3, the stylo-pharyngeal muscles; *c*, section of lower jaw; *d*, hyoid bone; *e*, thyroid cartilage; *g*, isthmus of thyroid gland; 4, cut edge of mylohyoid muscle; 5, 6, 7, 8, muscles of tongue; 9, 10, 11, superior, middle, and inferior constrictors of pharynx; 12, œsophagus. (Allen Thomson.)

When a mouthful of water is taken, two sounds may be heard on auscultating over the œsophagus. The first sound immediately follows the beginning of the act of swallowing and is probably due to the impact of the fluid against the posterior pharyngeal wall, brought about by the sudden contraction of the muscles which throw the fluid from the back of the tongue across the pharynx. The second sound is heard best by listening over the epigastrium. Since a certain amount of air is always swallowed at the same time as the food, the escape of this air through the small cardiac orifice gives rise to a bubbling noise which can be easily heard. It begins from four to ten seconds after the first sound, and lasts for two or three seconds. The interval between the two sounds is not constant. If the observation be carried out on a man lying on his back, the trickling second sound is changed into a series of sounds which have been described as squirts, which vary from two to five in number, each lasting about one second. The second sound may be absent when a solid bolus is swallowed.

On observing the process by X-rays, very much the same time relations are obtained. If a mouthful of milk mixed with barium sulphate be swallowed, it will be seen passing rapidly down the œsophagus to the cardiac orifice of the stomach. Here the passage is arrested for a time, and the fluid enters slowly in a narrow stream into the stomach. The average time which elapses between the beginning of deglutition and the disappearance of the last trace of fluid from the œsophagus is about six seconds (Fig. 553). A similar course of events is induced when the food swallowed is semi-solid. The bolus is shot at a surprising speed through the pharynx and upper œsophagus, but is slowed down at about the level of the clavicle, and travels more and more slowly till it reaches the cardiac orifice. If, however, the bolus be dry, such as a cachet, it may take as much as fifteen minutes to reach the cardiac orifice, although the individual who has swallowed it is quite unaware of its continued presence in the œsophagus. If, as would normally be the case, the cachet be well moistened with saliva or water before swallowing, it passes much more rapidly.

---

* Couvelaire and Crouzon, *J. de Physiol. et de Pathol. générale*, 1900, **2**, 280; Patel and Pellande, *J. de Physiol. et de Pathol. générale*, 1903, **5**, 47.

It has been customary since the time of Magendie to divide the act of swallowing into three stages: during the first, the bolus of food is carried past the anterior pillars of the fauces; during the second, through the pharynx, past the openings of the nasal cavities and of the larynx; and during the third, through the œsophagus into the stomach. There is, however, no pause between these various stages. The act of deglutition is one, and the initiation of the first stage inevitably involves the completion of the whole. The food when masticated is collected as a bolus on the dorsum of the tongue. A pause then takes place in the movements of mastication, and a slight movement of the diaphragm usually occurs known as "respiration of swallowing." An early event in swallowing is the closure of the mouth and nasal passages and slight raising of the larynx. If the thyroid cartilage is held down, or if the mouth is open to the pharynx, normal swallowing cannot occur. Next, a sudden elevation of the tongue throws the bolus back through the anterior pillars of the fauces. In this movement the chief factor is the contraction of the mylohyoid muscle, which presses the tongue against the palate and pushes it backwards. The backward movement of the tongue may also be aided by the contraction of the styloglossus and palatoglossus muscles, which pull the base of the tongue suddenly backwards. These muscles, especially the palatoglossi, serve to close the isthmus faucium, thus preventing any return of the food towards the mouth.

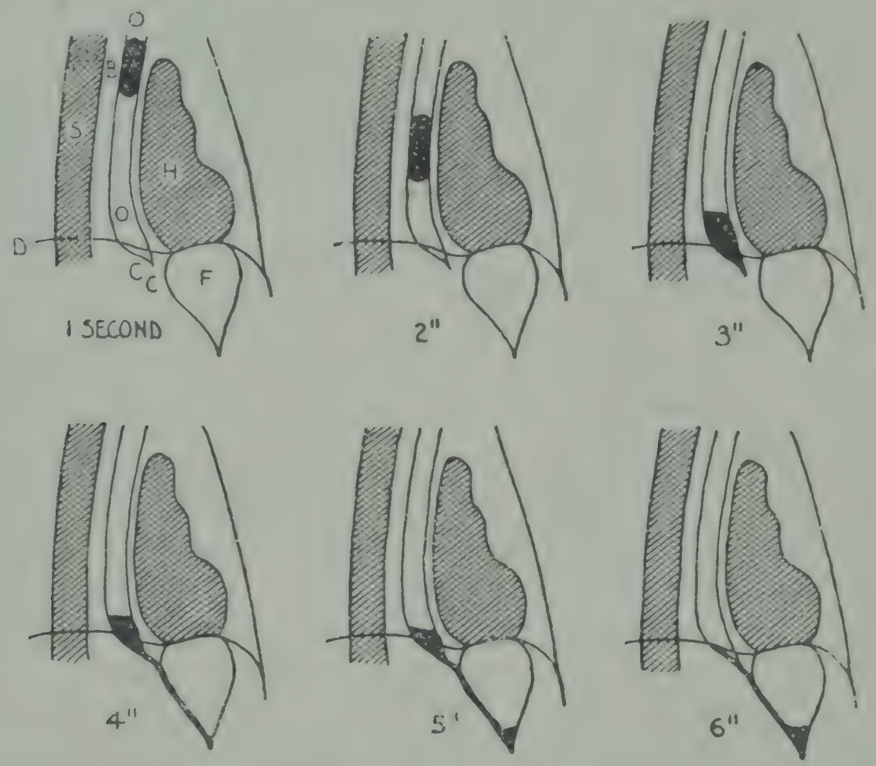


Fig. 553. Appearance of lower end of Œsophagus and Cardiac Sphincter, at one to six seconds after swallowing; observed under X-rays. (Hurst.)

As the food is passing through the upper part of the pharynx, it traverses a region common to the respiratory as well as the digestive passages. Its passage through this region is therefore rapid, and is associated with a closure of the two openings of the air passages into the pharynx. The nasal cavity is shut off by a simultaneous contraction of the levator palati and palato-pharyngeal muscles and azygos uvulæ, by which means the soft palate is raised (Fig. 554) and the posterior pillars are approximated to the uvula. The upper and back wall of the palate is thus formed into a tense sloping roof.

More important is the shutting off of the lower air passages. The contraction of the mylohyoid muscles, which initiates deglutition, is followed almost immediately (at an interval of 0·1 sec.) by an abrupt further elevation of the larynx, and this elevation is accompanied by closure of the glottis as well as of the superior opening of the larynx. The laryngeal opening is bounded in front by the epiglottis, behind by the tips of the arytenoid cartilages, and at the sides by the aryteno-epiglottidean folds. When deglutition takes place, the arytenoid cartilages, which normally lie against the posterior wall of the pharynx, are rotated and move inwards and forwards, so that the laryngeal opening assumes the form of a tri-radiate fissure, the vertical limb being short, while the transverse limb is rounded owing to the pulling inwards of the margins of the epiglottis. The epiglottis does not descend as a lid.* At the same time both the true and false vocal cords come together, while the movement of the dorsum of the tongue backwards enables the closed laryngeal orifice to lie directly under the back part of the tongue. The muscles which are actively involved in this closure of the lower air passages are the external thyro-arytenoid arytenoid, ary-epiglottidean, and the lateral crico-arytenoid muscles. Since the approximation of the posterior to the anterior boundary of the laryngeal opening is rendered possible only by the elevation of the whole larynx under the hyoid bone, the act of deglutition cannot be carried out unless the larynx is free to move.

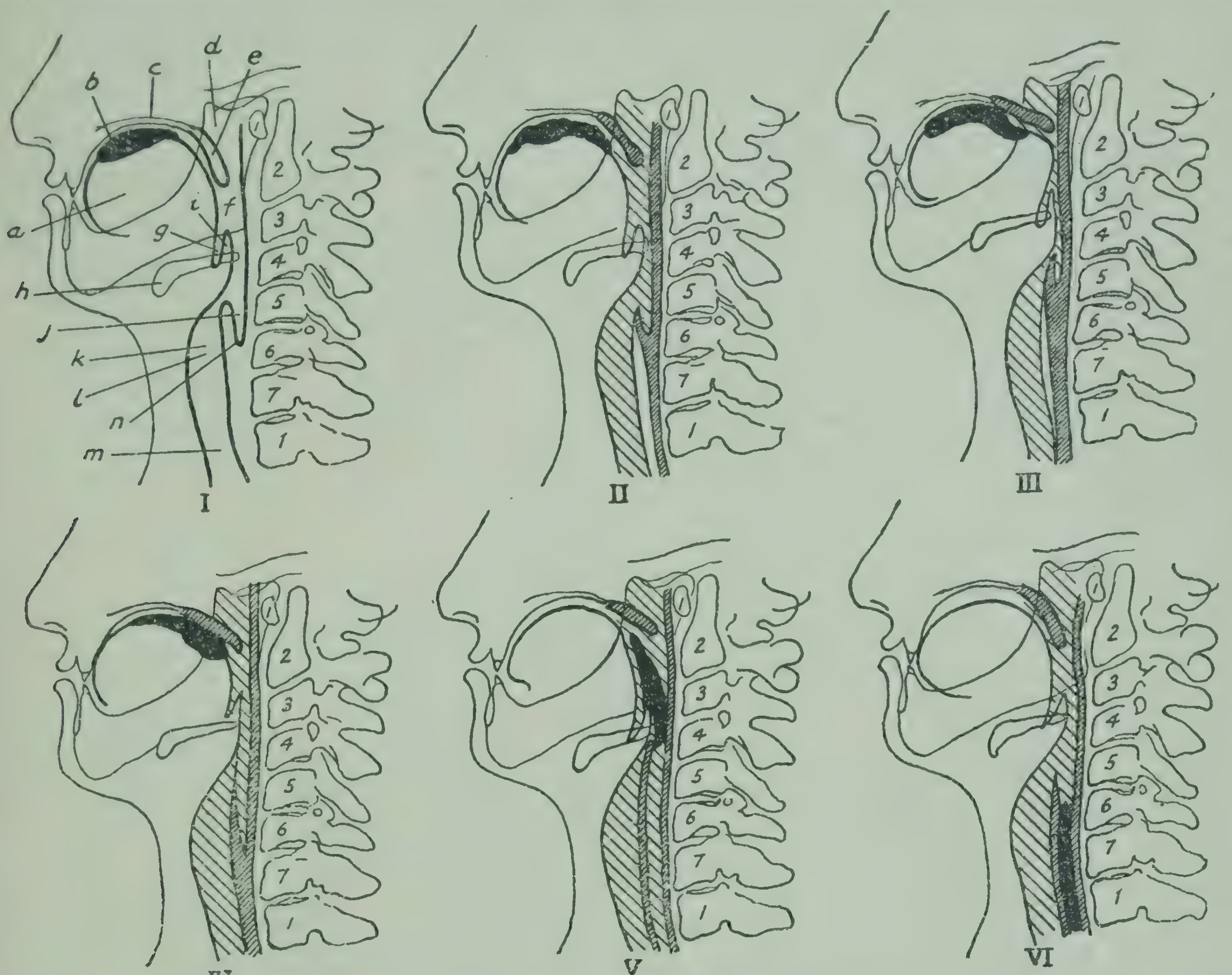


Fig. 554. Diagrams showing events in deglutition, as deduced from radioscopic observations. I. Before commencement: *a*, tongue, *b*, bolus, *c*, hard palate, *d*, nasopharynx, *e*, soft palate, *f*, pharynx, *g*, vallecula, *h*, hyoid, *i*, epiglottis, *j*, laryngeal pharynx, *m*, trachea, and *n*, œsophagus.
II. 1st stage, soft palate closing, larynx rising and pharynx narrowed.
III. 2nd stage: nose closed off, larynx raised, pharynx and upper larynx obliterated.
IV. Pharynx opening up, laryngeal pharynx drawn up to epiglottic position, œsophagus and upper larynx open.
V. Pharynx open, drawing bolus in.
VI. Laryngeal pharynx and œsophagus drop back with contents to normal position. (Barclay.)

The two openings from the back of the pharynx into the air passages being thus closed, the bolus is shot rapidly past them into the region of the middle and inferior constrictors of the pharynx. If the bolus be liquid, the movement of the back part of the tongue may be sufficient to propel the substance past the constrictors through the lax œsophagus to its lower end. It is on this account that, when corrosive fluids are swallowed, we very often find the damage to the œsophagus limited to the three points where it is narrowed and there is a slight hindrance to the onward flow of fluid. If the bolus be large and semi-solid, it nevertheless passes through the pharynx with great speed. One view of what happens here is that it is seized in the grasp of the middle constrictors on passing through the upper part of the pharynx, and is thrust by successive contractions of this muscle, and of the inferior constrictor, into the œsophagus. The walls of the cervical part of the œsophagus are composed of striated muscle. In the thorax, striated and unstriated muscles are associated together, while the lower third, in the neighbourhood of the stomach, consists almost entirely of unstriated muscle. Corresponding to these differences in structure, Kronecker and Meltzer * found differences in the duration and rapidity of propulsion of the contraction waves in each part. The contraction of the constrictors must be very rapid if this explanation is correct. Registration of pressures at different levels has not always given the same results in different investigations. According to Kronecker and Meltzer, a series of positive pressure changes travels along pharynx and œsophagus. Others have

---

* Anderson Stuart. *J. Physiol.*, 1906, **35**, 446.

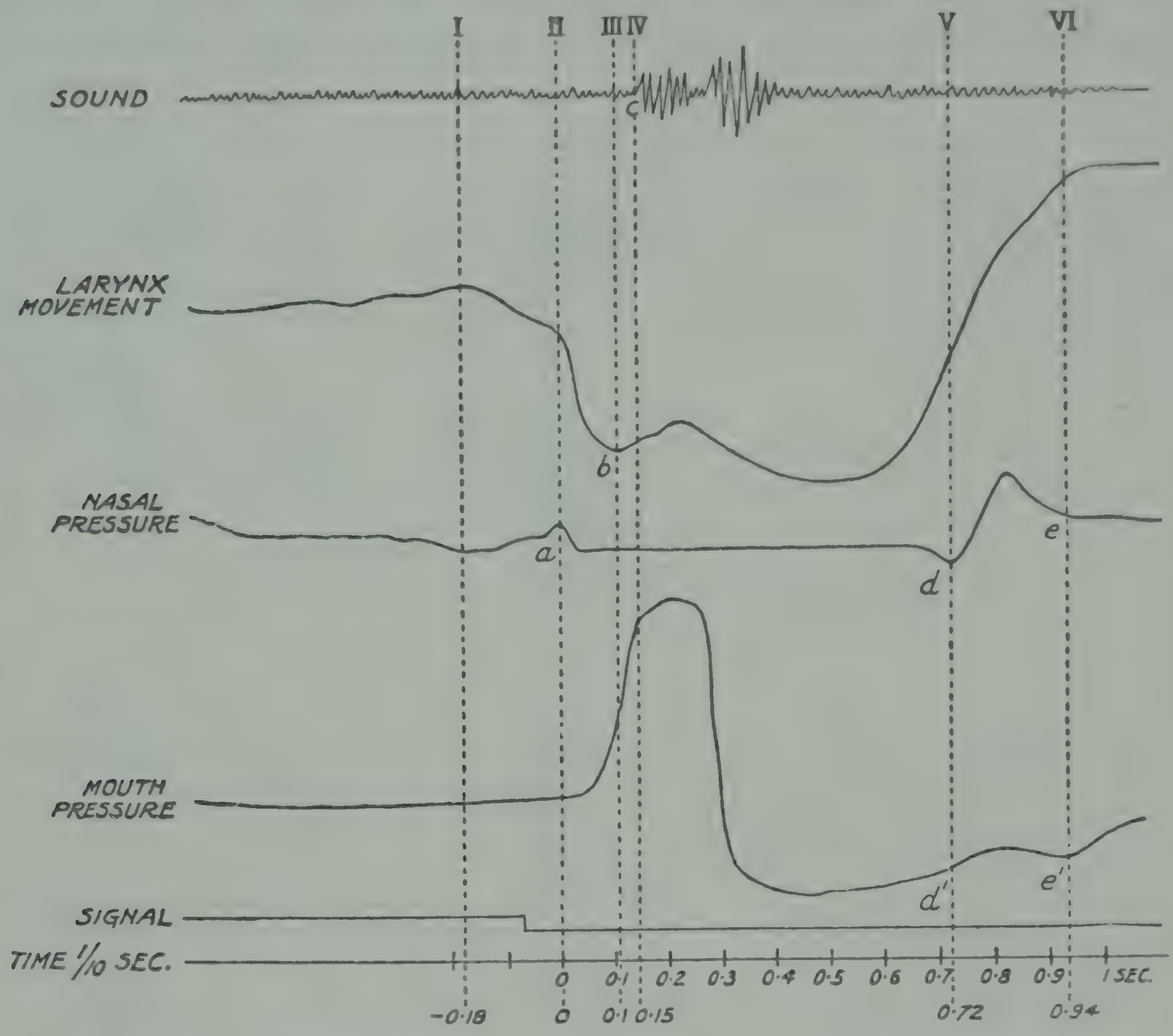


Fig. 555. Records during swallowing. In order from above down.
Record of first sound (string galvanometer and microphone).
Movement of thyroid cartilage.
Pressure in nares.
Pressure in mouth.
Signal by subject showing commencement of swallowing.
Time in 0·1 sec.

The vertical lines show: I. Commencement of voluntary part of act: II. Nasal cavity shut off at (*a*); commencement of contraction of mylohyoid and of reflex part of act. III. Movement of larynx at (*b*); contraction of geniohyoid, genioglossus and hyoglossus. IV. Sound at (*c*). V. Opening of nasal cavity at (*d*), followed by rise of pressure. VI. End of act. (Kaiser, *Arch. néerl de Physiol.*, 1918, I. 148.)

---

* Kronecker and Meltzer. *Arch. f. Anat. u. Phys.* 1880, 299, 446; 1883, *Suppl.* 328.

obtained negative pressures in the pharynx and upper œsophagus (*vide* Fig. 555) preceding the positive wave.* This would agree with the radioscopic observations of Barclay,† who describes the following events :

" (1) The mouth and nose are closed.

(2) The larynx is raised and closed and its upper part is obliterated by the back of the tongue.

(3) The pharyngeal space is obliterated by (*a*) raising the larynx, and (*b*) retraction of the tongue. (The posterior wall of the pharynx does not come forward ; the pharynx is not constricted but is compressed from before backward against the posterior wall, that is, against the vertebræ.)

(4) The pharynx is opened up, and with the three outlets closed a negative pressure is obtained by (*a*) dropping the larynx, and (*b*) allowing the tongue to go forward.

(5) During this movement the laryngeal pharynx still remains high, in close relation to the epiglottis. It is in an open-mouthed condition.

(6) Food is tipped over the back of the tongue.

(7) Food is sucked into the open mouth of the laryngeal pharynx, which, as it receives the bolus, drops from the epiglottis down to its normal position, thus opening up the larynx again, after the food has passed.

(8) The food is probably sucked some distance down the œsophagus."

" The mechanism by which the larynx is cut off during the act is by a dragging up of the laryngeal pharynx behind the larynx to form contact with the epiglottis, possibly about a quarter of an inch below the tip of the epiglottis."

The free passage of liquids down the œsophagus under the influence of the propulsive force exercised by the mylohyoid muscles shows that the walls of this tube must be lax, and in fact one must assume that the first act of deglutition, so far as concerns the œsophagus, is an inhibition initiated reflexly with the beginning of the act of deglutition. When a second act of deglutition succeeds the first with a sufficiently short interval, the reflex inhibition due to the second act may prevent the development of any wave of contraction in the œsophagus. This tube thus remains in a lax condition and allows the free rapid passage of the food downwards until the movements of deglutition have come to an end, when a peristaltic contraction occurs and sweeps all remaining adherent particles of food into the stomach.

The circular fibres of the lower end of the œsophagus, which form the cardiac sphincter of the stomach, are normally in a state of tonic contraction. When one mouthful of food is swallowed, it may either pass directly into the stomach by gravity, or remain at the lower end of the œsophagus until the following peristaltic wave forces it through the orifice. Entry into the stomach often occurs as an expiration begins. Distension of the œsophagus causes peristaltic œsophageal contractions (Fig. 556) of simpler form than those resulting from conscious swallowing ; these contractions move the bolus forwards if it tends to stick. Considerable distension causes pain beneath the sternum.‡ When several acts of deglutition succeed one another, the cardiac sphincter shares in the in-

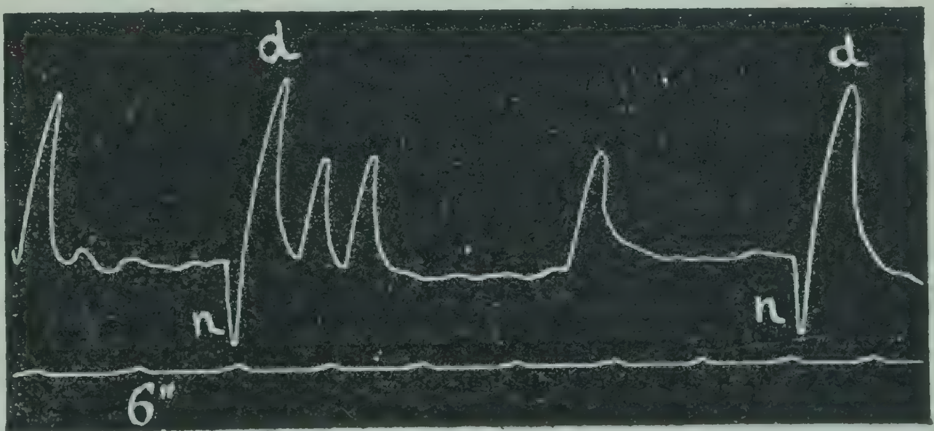


Fig. 556. Contractions of œsophagus recorded in man by a balloon method. At *n*, swallowing movements occurred. The subject was unaware of the other contractions. Note the negative wave as first event in conscious swallowing. (Danielopolu *et al.*)

---

* Schreiber, *Arch. f. exp. Path. u. Pharm.*, 1901, **46**, 414 ; **67**, 72 ; Danielopolu, Simici and Dimitriu, *J. de Physiol. et de Path.*, 1924, **22**, 595.

† Barclay. *Brit. Journ. Radiol.*, 1930, **3**, 535 ; *Acta Radiol.*, 1932, **13**, 91.

‡ Poulton and Payne. *J. Physiol.*, 1922, **56**, Proc. liii.

hibition of the œsophageal walls, and offers no resistance to the direct propulsion of food from the mouth to the stomach. Cannon has shown that the relaxation of the cardiac orifice which accompanies swallowing, extends also to the cardiac end of the stomach. This relaxation lowers the pressure within the stomach, and makes room for the incoming food.

When the contents of the stomach are only feebly acid, the cardiac sphincter opens at intervals, especially if the pressure in the stomach is high, and allows of the regurgitation of the stomach contents into the lower part of the œsophagus. This is at once followed by a peristaltic contraction of this part of the œsophagus (apparently entirely unconscious), which drives the fluid back into the stomach. It has been supposed that movements of regurgitation become more and more infrequent as the gastric contents become acid ; but the modern view associates the increasing tone of the sphincter with distension of the stomach and the stage of digestion.

The Nervous Mechanism of Deglutition. Deglutition is a reflex act. Swallowing reflexes are readily elicited in the decerebrate animal, when water or dilute alcohol is placed in the mouth. When we swallow voluntarily we supply the necessary initial stimulus, either by touching the fauces with the tongue, or by forcing saliva or food into the fauces. Normal swallowing is abolished by cocainisation of the mucosa of the pharynx. The afferent channels of the reflex are contained in the second division of the fifth nerve, the glosso-pharyngeal nerve, and the pharyngeal branches of the superior laryngeal nerve. We can excite a single act or a whole series of acts of deglutition by electrical stimulation of the central end of the last-named nerve. The efferent fibres travel by the hypoglossal nerve to the muscles of the tongue, by the fifth to the mylohyoid, by the glossopharyngeal, the vagus and the spinal accessory nerves to the muscles of the fauces and pharynx. The closure of the larynx is effected by impulses which travel through the superior and inferior laryngeal branches of the vagus. Normal swallowing is impossible after section of both nerves near their origins. The centre for the act is situated in the medulla oblongata, and can be considered as consisting of a chain of centres, stimulation of one of which involves the firing off of all the others in orderly sequence ; the propagation of the contraction down the œsophagus is determined by the intracentral nervous connections, and does not require the integrity of the muscular tube itself. If the œsophageal nerves be divided, the act of deglutition is abolished, the upper part of the œsophagus becoming permanently relaxed, while the lower part,

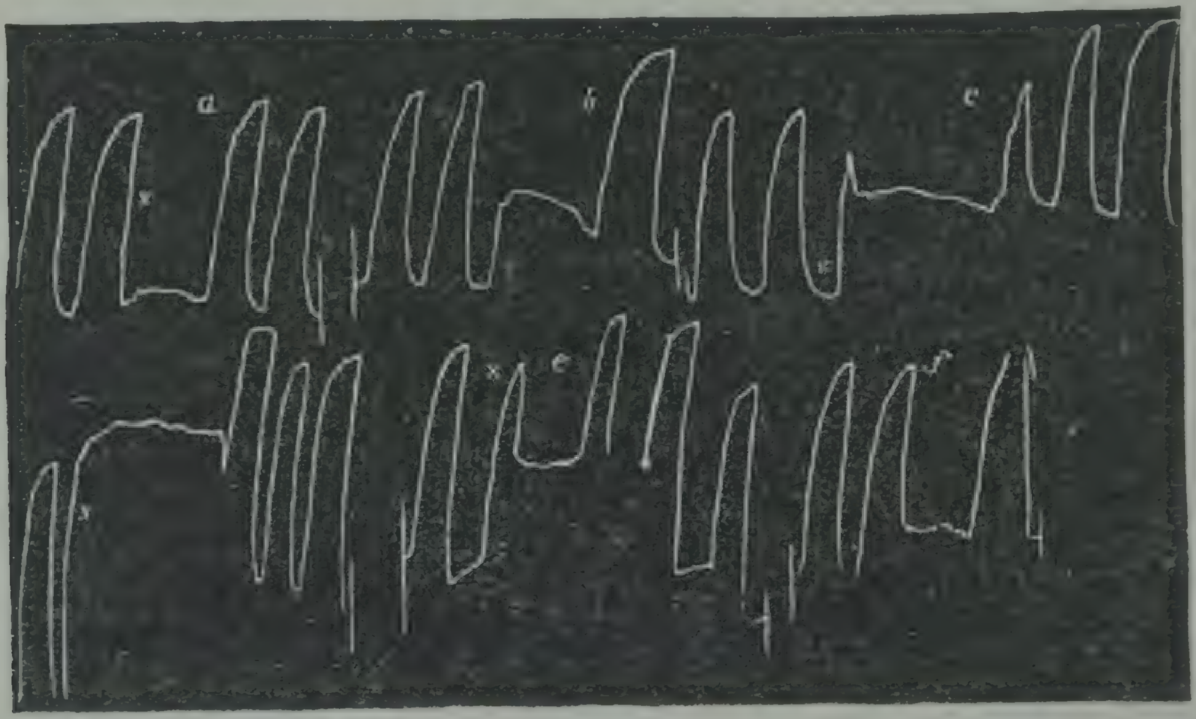


Fig. 557. Tracings of Respiratory Movements to show the Effect of Stimulating the Central End of the Glossopharyngeal Nerve. (Marckwald.)

The point of stimulation is marked with a cross. Note that the stoppage may occur at any phase of the respiratory movement.

including the cardiac sphincter, enters into a state of tonic contraction. On the other hand, the œsophagus may be ligatured or cut across without interfering with the progression of the wave of contraction, started in the pharynx, from one end of the tube to the other. Stimulation applied to the mucous surface of the œsophagus is without effect.

There is an important interdependence between the functions of respiration and deglutition. If an inspiratory or expiratory movement were going on during the act of deglutition, food might be drawn into the lungs or driven into the nasal cavities. Such an accident is prevented by the fact that every act of swallowing inhibits a respiratory movement. This can be shown on oneself. If the breath be held until the feeling of dyspnœa becomes insistent, relief is at once experienced by swallowing, and the feeling of relief will last for three or four seconds. In some animals this inhibition appears to be effected reflexly through the glossopharyngeal nerve. Stimulation of the central end of this nerve in the rabbit * at once causes cessation of respiration in whatever phase it may happen to be (Fig. 557). This cessation lasts for five or six seconds, *i.e.* a sufficient length of time for a whole series of acts of deglutition. Respiration then recommences, and the inhibition cannot be prolonged by continuing the stimulation of the glossopharyngeal nerve. In other animals, *e.g.* cat, there may be acceleration of breathing, while the dog gives variable results.† The afferent path is thus doubtful.

3. DIGESTION IN THE STOMACH

The Gastric Contents. The normal stomach several hours after food is rarely empty, but usually contains from 30 to 50 c.c. of fluid consisting of gastric juice, mucus, saliva and regurgitated duodenal contents. Within five minutes of the taking of food into the mouth, a secretion of gastric juice begins from the multitude of tubular glands which make up the greater part of the mucous membrane of the stomach. As the food is swallowed in successive portions, it accumulates in a mass in the body of the stomach, and the mass thus formed is penetrated with difficulty by the gastric juice,‡ so that salivary digestion can be continued for a considerable time. The flow of juice may continue for several hours and amounts in man, after a good meal, to 500 c.c. or more.

Gastric contents for examination can be obtained in various ways. The present method (fractional test meal) consists, after administering a standard meal (*e.g.* gruel), in withdrawing small samples of the contents at 15-minute intervals through a narrow tube which can be swallowed until its lower opening is about 20 inches from the incisor teeth, and left in place throughout the examination (Rehfuss).

Gastric Juice. The earliest case in which human gastric secretion was adequately studied was in a man with a permanent gastric fistula.§ Similar cases are studied from time to time.‖

Human gastric juice can best be obtained by the subcutaneous injection of 0·5 mg. histamine hydrochloride (or 0·1 mg. per 10 kg.), the stomach having been previously washed out. In normal subjects this leads to a secretion of gastric juice at the rate of about 200 c.c. per hour : the juice as secreted probably contains 0·5 to 0·6 per cent. HCl, but this is reduced by neutralisation by mucus, which is also secreted throughout. The free acid of the juice varies considerably in consequence, the average in healthy men being about 0·39 per cent. HCl (= 107 c.c. n/10 per 100 c.c. ; standard deviation 31·5 ; standard error 2·42) ¶ ;

---

* Marckwald. *Z. f. Biol.*, 1887, **23**, 149, 239.

† Teitelbaum and Ries. *Amer. J. Physiol.*, 1935, **112**, 684. (Various early workers also found acceleration.)

‡ Grützner. *Pflüger's Arch.*, 1905, **106**, 463.

§ William Beaumont, " Experiments and Observations on the Gastric Juice and the Physiology of Digestion," Plattsburgh, 1833.

‖ Wolf and Wolf. *Amer. J. Physiol.*, 1943, **138**, 309.

¶ Lee Lander and Maclagan. *Lancet*, 1934, **227**, 1210.

more individuals are above the mean than below it. The juice contains less pepsin in the later than in the earlier stages of secretion.*

It is sometimes necessary in a patient to perform gastrostomy, *i.e.* to make into the stomach an artificial opening through which he can be fed. In such patients it is observed that merely taking food into the mouth is followed by the secretion of gastric juice. Pavlov produced this condition artificially in dogs. The œsophagus was divided and the two ends brought to the surface of the neck. At the same time an opening was made into the stomach (Fig. 558).

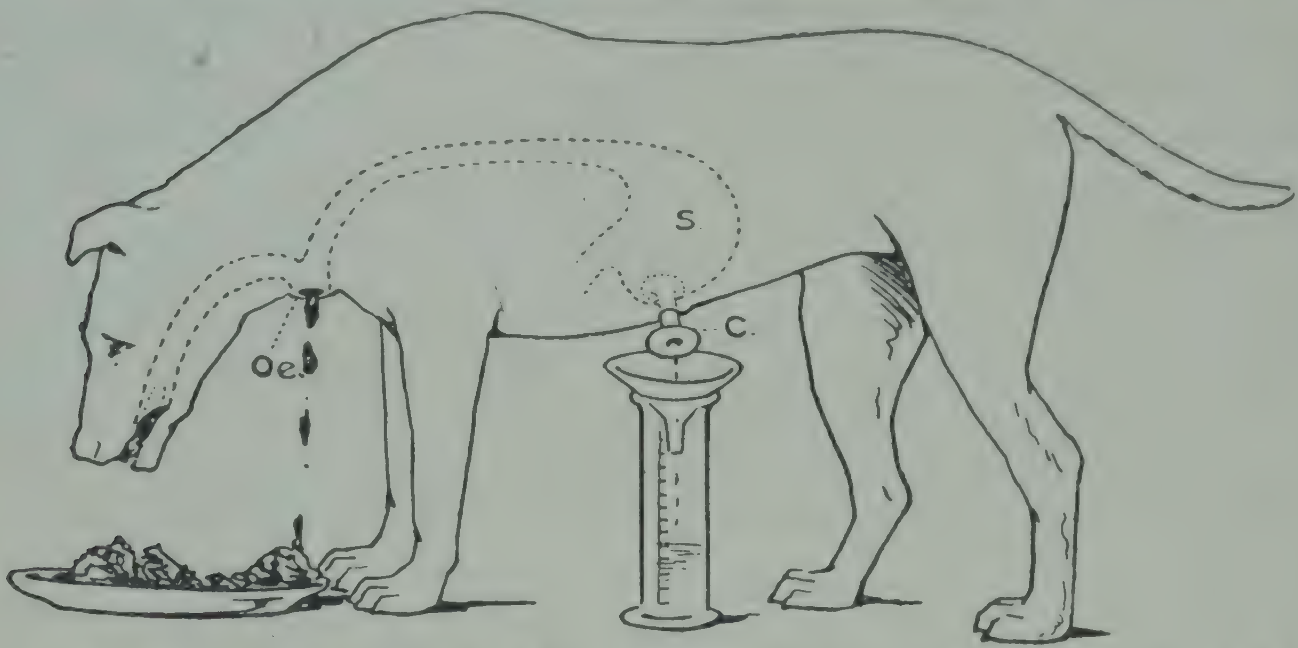


FIG. 558. Illustrating Pavlov's Method of obtaining Pure Gastric Juice. Oe, the œsophagostomy opening ; S, the stomach, with fistula in which is fixed the silver cannula C.

The animals could eat and swallow food, but the food thus swallowed fell out of the fistula in the neck instead of passing into the stomach. Under these circumstances it is found that the "sham feeding" is quickly followed by a secretion of gastric juice. If taken from a fasting animal, such a juice can be regarded as pure gastric juice. It is clear, colourless, strongly acid (pH about 1·0), and without smell. It contains only traces of protein. The following table represents its average composition.†

Hydrochloric acid	0·5 p.c. (= 136 m. eq./l.)
Other Chloride	0·107 p.c. (= 30 m. eq./l.)
Na	0·05 p.c. (= 22 m. eq./l.)
K	0·028 p.c. (= 7·2 m. eq./l.)
Ca	0·02 p.c. (= 0·5 m. eq./l.)
Total solids	0·43 to 0·65 per cent.
Enzymes and other organic solids	— ,, 0·45 ,,
Total nitrogen	0·05 ,, 0·08 ,,

Freezing point = − 0·59° C.

The actions of gastric juice are due partly to the acid, mainly to the combined action of the acid and the enzymes.

The free acid of the gastric juice is entirely hydrochloric acid. Dog's juice contains on the average about 0·5 per cent. HCl. When, however, we examine the gastric contents, composed of a mixture of gastric juice and semi-digested food, we also often find lactic acid which is produced by fermentation occurring in the carbohydrates. As the gastric juice gradually renders the contents acid, it stops this lactic acid fermentation. The free H ions present, as measured by indicators, or by the hydrogen electrode, are diminished considerably by the presence of proteins, proteoses or peptones, which act as weak buffers. But it does not seem that the adjuvant action of the acid on the proteolytic powers of the gastric juice is in any way affected by the diminution of its acidity ("combined HCl") caused by the presence of peptone.

In some pathological conditions, *e.g.* pernicious anaemia and gastric carcinoma, free hydrochloric acid may be entirely wanting from the gastric juice and the detection of this acid in gastric contents becomes therefore a matter of considerable clinical importance. For this purpose we can employ various indicators, which change colour in the presence of a strong acid such as HCl, but are unaffected by weak acids such as lactic acid or the fatty acids. The reagent which is most employed is Günzberg's reagent, a mixture of phloroglucinol and vanillin in absolute alcohol. A drop of this is evaporated to dryness. A drop of the fluid to be tested is then added, and also evaporated to dryness. If free HCl be present, the residue becomes a brilliant red colour, an effect which is not produced by the presence of free organic acids.

---

* HELMER. *Amer. J. Physiol.*, 1934, **110**, 28.

† Based in part on figures by GRAY and BUCHER, *Amer. J. Physiol.*, 1941, **133**, 542.

In order to determine quantitatively the amount of total acidity and free HCl, the gastric content is titrated with $N/10$ NaOH using thymol blue as indicator The reading is taken when the red colour changes to yellow-orange, and gives the free HCl. The titration is then continued until the indicator turns blue ; the further amount of alkali needed for this represents the combined HCl and organic and other weak acids. Probably the most accurate method is to titrate with standard NaOH after adding to the juice nine times its volume of acetone, using naphthylamine orange as indicator : end points are at pH 3·4 (pink) and pH 4·0 (orange). The weak acids are so much further weakened by the acetone that they scarcely require any alkali as compared with free HCl.* Estimation of total chlorides and mineral chlorides is also valuable.†

The Actions of Gastric Juice

Acid. By the action of the hydrochloric acid, certain changes are induced in the foodstuffs. Cane sugar is inverted to glucose and fructose ; a small amount of hydrolysis also takes place in the dextrins and maltose produced by the action of ptyalin on starch. Some proteins, such as blood fibrin, and the collagen of the connective tissues, are swollen up to form a jelly-like mass.

Pepsin. The chief digestive function of the gastric juice is dependent on the action of the enzyme pepsin. (Cathepsin is present also.‡ It is a proteolytic enzyme functioning at pH 4–5 : its maximal activity is obtained in presence of certain activators, *e.g.* H_2S, cysteine, etc.)

Pepsin, which is inactive in neutral medium, is actively proteolytic in strongly acid solution (optimum of pH 1·5–2·0), hydrochloric acid being the natural acid. Its effect is on the proteins, and may be best studied on blood fibrin. If fibrin be immersed in 0·4 per cent. hydrochloric acid, it swells up to a gelatinous mass. On adding chopped gastric mucosa, or pepsin, and warming to 37° C., the gelatinous mass rapidly undergoes solution. After some hours, a whole series of products of hydrolysis is found in the mixture. On boiling and neutralising the fluid, a precipitate may be thrown down consisting chiefly of *acid-meta-protein.* After boiling and filtering the filtrate contains *proteoses* and *peptones.* The proteoses are precipitated by saturation with ammonium sulphate, by which peptones are not precipitated.

In the course of peptic digestion it can be shown that COOH groups and NH_2 groups become liberated, and in equal numbers,§ which show that pepsin acts by breaking certain of the peptide linkages of proteins, thus :—

$$\cdots R.CO.NH.R'\cdots + H_2O = \cdots R.COOH + NH_2.R'\cdots$$

After completion of its action, about one-tenth of the total peptide linkages of the protein molecule have been severed in this way.

Proceeding from primary through secondary proteoses to peptones, there is a continuous diminution in the size of the molecule. During the time over which gastric juice can act, a maximum, say, of eight hours, the breakdown never passes beyond the proteose and peptone stage, and it is in this form, and as acid meta-protein, that the proteins of the food pass on into the small intestine. Analysis of gastric contents in man shows that, whereas carbohydrates may have been largely converted into reducing sugar in an hour, less than 3 per cent. of the protein nitrogen had been resolved into non-protein derivatives.||

* Richardson. *Proc. Roy. Soc.*, 1934, **115B**, 170.

† *v.* Harrison. " Chemical Methods in Clinical Medicine," 1947. (Churchill.)

‡ Taylor *et al.* *J. Clin. Invest.*, 1938, **17**, 335 ; Freudenberg, *Enzymologie*, 1940, **8**, 385.

§ Waldschmidt-Leitz and Künstner. *Z. f. Phys. Chem.*, 1927, **171**, 70.

|| Beazell. *Amer. J. Physiol.*, 1941, **132**, 42.

Action on particular Proteins. *Collagen.* The connective tissues are made up chiefly of collagen, which on prolonged boiling is converted into gelatin. The gastric juice dissolves collagen, converting it into gelatoses and gelatin peptones. On account of this action, adipose tissue is broken up. The cellular envelopes are dissolved, and the fat floats freely in the gastric juice.

Elastin, the chief constituent of elastic fibres, is indigestible.

Mucin, which forms a considerable proportion of the ground substance of connective tissues, is converted by gastric juice into peptone-like substances, and into reducing bodies allied to glucosamine.

The *Nucleo-proteins* are broken up into two moieties. The protein half is converted into proteoses and peptones, the nuclein half into *nucleic acid.*

CHYMOSIN. Caseinogen of milk is present in the form of calcium *caseinogenate.** The first effect of gastric juice, even in neutral medium, is to convert the caseinogen into an insoluble form called *casein.* This action is often in part due to the presence of a distinct enzyme named *chymosin* or *rennin.*† Although crystalline pepsin also clots milk, an active chymosin preparation can be obtained which has no peptic activity, is a thioproteose and in various other respects differs chemically from pepsin.‡ It has been stated that adult human gastric juice contains no chymosin.§ For the clotting of milk by chymosin the presence of calcium ions is necessary. The addition of chymosin to oxalated milk apparently produces no effect, but clotting ensues if calcium chloride is then added to the mixture. It is therefore thought that, in the clotting of milk, the caseinogen first undergoes conversion into casein, the sodium salt of which is soluble, but the calcium salt insoluble. Hence if calcium ions are present, calcium caseinate, or insoluble casein, is precipitated. Under the further action of the gastric juice, the clot of casein first formed is dissolved.

Action on Carbohydrates. The chief digestion of carbohydrate which takes place in the stomach is that of starch, since the amylase of the swallowed saliva continues to act until the contents become definitely acid. The hydrolysis of cane sugar and some other carbohydrates, *e.g.* inulin, which takes place in the stomach can be accounted for by the action of hydrochloric acid.

Action on Fats. The chief action of this juice on fats is the solution of their connective tissue framework and protoplasmic envelopes, so as to set the fat free in the gastric contents. After a fatty meal it is found, moreover, that a proportion of the fat in the stomach has undergone hydrolysis with liberation of free fatty acid. In this hydrolysis three factors are involved : (1) the action of the warm dilute hydrochloric acid ; (2) the action of a *lipase,* which is secreted by the walls of the stomach, and acts especially at the beginning of gastric digestion before the contents have attained a high degree of acidity. The action of this enzyme is conspicuous only if the fat be present in a finely divided form, as in yolk of egg ; (3) the action of lipase regurgitated from the duodenum. In any case, the chief digestion of fat takes place in the duodenum.

ANTI-ANÆMIC FACTOR. Gastric juice also contains an important principle called the *intrinsic factor* which reacts with substances present in certain foods to form a substance essential to the proper maturation of red blood corpuscles (p. 520). The intrinsic factor is thermolabile.

The Secretion of Gastric Juice

Source of Components. The stomach consists of the *cardiac portion,* the *corpus* or main body, and the *pyloric part.* There are definite differences between these as regards their functions.

---

* In the U.S.A. and on the Continent caseinogen and casein are respectively called *casein* and *paracasein.*

† In French, *présure* ; in German, *Lab* ; Italian, *presame.* The English name, *rennin,* is derived from "rennet-bag," the fourth stomach of the calf. An International name such as *chymase* or *chymosin,* is clearly preferable.

‡ TAUBER and KLEINER. *J. Biol. Chem.,* 1932, **96,** 745 ; 1934, **104,** 259.

§ DOTTI and KLEINER. *Amer. J. Physiol.,* 1943, **138,** 557.

Four types of cells are recognisable in the gastric mucosa, viz. surface, mucoid, peptic and oxyntic cells.* The surface cells cover the inner surface of the mucosa, and line the ducts of the tubular gastric glands; they secrete mucus. Mucoid cells form the secreting cells of those pyloric and cardiac glands which are near the two sphincters; they are also mingled with the peptic cells lining the cardiac glands. They are probably connected in function with the peptic cells, since in new-born animals they precede these in the cardiac glands.

Oxyntic, or parietal, cells are absent from the pyloric part of the mucosa, and the juice secreted in the pyloric portion contains no free HCl. A survey of the distribution of enzymes in sections of gastric mucosa, when correlated with the distribution of the various cells, shows that pepsin is related to peptic cells and a peptidase to the peptic and neck cells.†

The pepsin is not present as such in the resting peptic cells, but is represented by a zymogen, *pepsinogen*, present in the form of granules. ‡ As secretion proceeds the granules diminish in number, and are confined to the inner margin of the cell. Pepsinogen has been isolated in pure crystalline form from the mucosa of the body of the stomach.§ It is converted autocatalytically into pepsin in solutions more acid than pH 6, and in the process about 15 per cent. of the nitrogen content is split off as a polypeptide.

The secretion of a strong acid in concentrations of about one-tenth normal must be admitted to be an arresting phenomenon. It has long been known that the secretion of acid is effected by the oxyntic cells; it is also known that these contain large amounts of the enzyme carbonic anhydrase. Several theories have been put forward to explain the secretion. Ultimately the H-ion is derived from water:—

$$H_2O \longrightarrow H^{\cdot} + OH'$$

and the energy for separating the ions of water and transporting the $H^{\cdot}$-ion towards the lumen of the gland must be provided by the aerobic metabolism of the oxyntic cells.

The OH'-ion has to be neutralised and this is done by uptake of CO_2, from the blood, and this, under the action of carbonic anhydrase, goes to H_2CO_3, which then reacts to give the HCO_3' ion ‖ :—

$$CO_2 + H_2O \rightleftharpoons H_2CO_3$$
$$H_2CO_3 + OH' \rightleftharpoons H_2O + HCO_3'.$$

This HCO_3' passes into the blood plasma. Electric neutrality is maintained by the entrance of Cl' ions which pass through the cell and are secreted with the $H^{\cdot}$ ions as HCl.

It has been shown that there is normally an electrical potential difference between the inner and outer surfaces of the frog's resting gastric mucosa, the inner side being about 20–30 mV. negative. When secretion occurs there is an increase in the rate of metabolism, and electrical energy is used to set up an electrochemical cycle *within* the cells, and this results in a diminution of the resting potential difference between the two sides of the mucosa. It is this electron-cycle which produces the separation of H^+ and OH' ions, and draws them to opposite sides of the cell. In support of this is the fact that if a current is passed through the mucosa from outer to secretory surface, *i.e.* enhancing the resting potential, HCl is secreted at an increased rate and if in the opposite direction the secretion is reduced or stopped.¶

Analogous explanations can be given for the secretion of alkali in the pancreatic juice and of acid in the urine.

There is an even higher concentration of carbonic anhydrase in the oxyntic cells than there is in red blood corpuscles; in the peptic cells there is very little.**

There certainly appears to be a close relationship between the total CO_2 content of the blood and the acidity of the gastric juice secreted.†† When dyes are injected into the blood, only those with an electropositive coloured ion are secreted into the gastric juice, those with electronegative chromogen ions being retained in the blood. This is due to the electric charge on the secreting membrane.‡‡

The normal resting stomach secretes some gastric juice, which appears to consist of a mucilaginous secretion, often rich in pepsin, but containing little or no acid. During and after a meal, brisk secretion of acid gastric juice occurs and persists for some time. Secretion as a result of mechanical stimulation of the gastric mucosa was described by Beaumont. Distension of the stomach readily causes secretion of acid juice.

---

* Lim. *Q. J. Micr. Sci.*, 1922, **66**, 187.

† Linderstrom-Lang *et al.* *C. R. Lab. Carlsberg*, 1935, **20**, i., 66.

‡ Langley. *J. Physiol.*, 1882, **3**, 269.

§ Herriott. *J. Gen. Physiol.*, 1938, **21**, 501.

‖ Davies and Roughton. *Biochem. J.*, 1948, **42**, 618.

¶ Crane, Davies and Longmuir. *Biochem. J.*, 1948, **43**, 321, 336.

** Davenport. *J. Physiol.*, 1939, **97**, 32.

†† Apperly and Crabtree. *J. Physiol.*, 1931, **73**, 331; Browne and Vineberg, *ibid.*, 1932, **75**, 345. Theorell. *Skand. Arch. f. Physiol.*, 1933, **66**, 225.

‡‡ Ingraham and Visscher. *J. Gen. Physiol.*, 1935, **18**, 695.

Secretory Nerve. The efferent gastric nerves are the vagus and the sympathetic. That the former is secretory is shown by direct stimulation of the vagus, which was carried out by Pavlov as follows: An animal with fistulæ of œsophagus and stomach had one vagus nerve divided in the neck. Four days later, the vagus was drawn out of the wound and its peripheral end stimulated. No effect was produced on the heart, owing to the degeneration of the cardio-inhibitory fibres. After five minutes stimulation the first drop of gastric juice appeared and a steady secretion was obtained with continuation of the stimulation. The length of the latent period before gastric secretion begins has not been satisfactorily explained. The vagus also controls secretion of all the mucous and mucoid cells of the stomach.* That it is the essential secretory nerve is proved by the fact that reflex secretion of gastric juice is abolished by section of both vagi. It also contains secretory inhibitory fibres, which have the effect of stopping or diminishing the secretion under certain conditions.

The sympathetic supply is vaso-constrictor, but stimulation of the splanchnic nerves five days after section (to allow degeneration of vaso-constrictors) gives a secretion of gastric juice, as also does the injection of adrenaline. This secretion is probably of subsidiary importance, and its significance is doubtful. In emotional states the secretion of gastric juice is greatly reduced.

Reflex Secretion. In "sham feeding", secretion of gastric juice begins in about five minutes after the onset of feeding, and continues as long as this goes on, which can be for hours. This method does not, however, enable us to determine whether the character of the juice will be altered in any way by the presence of food in the stomach itself. In order to form an idea of the normal course of secretion of gastric juice the "Pavlov pouch" † was devised. A small diverticulum, representing about one-eighth of the whole stomach, is made in the body of the stomach, in direct vascular and nervous continuity with the rest of the stomach, but shut off from the main part of the viscus by a double diaphragm of mucous membrane (Figs. 559, 560). In a dog treated in this way it is found that the amount of juice secreted by the small stomach is always a constant fraction of the amount secreted by the large stomach, while the digestive powers of the two juices are equal.

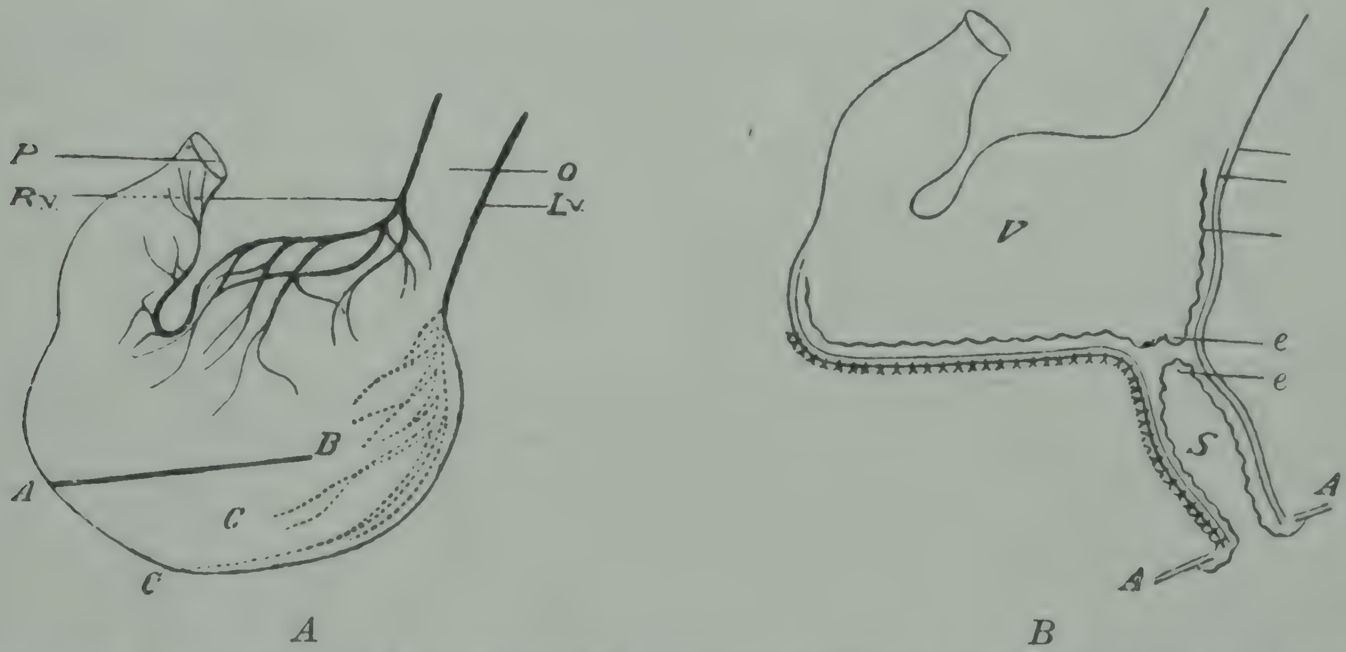


Fig. 559. Diagram to show Pavlov's Method of making a Cul-de-sac of the Body of the Stomach, with Vascular and Nerve Supply intact.

In A the line of the incision into the gastric wall is shown. B. represents the operation as completed.

In A: *O*, œsophagus: *R.v*, *L.v.*, right and left vagus nerves; *P*, pylorus *C*, corpus, or body of stomach; *A.B.* line of incision.

In B: *V*, main portion of stomach; *S*, the cul-de-sac; *A*, abdominal wall; *e*, *e*, mucous membrane reflected to form diaphragm between the two cavities.

We may therefore regard the secretion obtained from the stomach pouch as a sample of that produced by the rest of the organ, and from the juice collected from the pouch infer the behaviour of the whole stomach, *e.g.* the effects of a normal meal in which the food is swallowed, or of a sham meal in an animal with an œsophagostomy, or of a meal directly introduced into an opening into the large stomach.

---

* Jennings and Florey. *Q. J. Exp. Physiol.*, 1941, **30**, 329.

† The "Heidenhain" pouch is not innervated. Cope, *Arch. Surg.*, 1940, **40**, 717, has described a modified Pavlov pouch.

The fact that sham feeding causes the production of gastric juice shows that we have to do, in the first place, with a reflex nervous mechanism. This is an unconditioned reflex. But a secretion, which is at least as vigorous as that produced by a sham meal, can be evoked through a conditioned reflex, by merely providing some stimulus associated with the animal's food habits. If the animal be hungry, it is sufficient to show it the food to produce this effect. In the experiment from which the following Table is taken, the dog was continually excited by showing it meat during a period of an hour and a half. At the end of this time the animal, which had an œsophageal fistula, was given a sham meal. It will be observed that the "psychical secretion" obtained during the first period of the experiment was rather greater than the secretion produced by the introduction of food into the mouth.

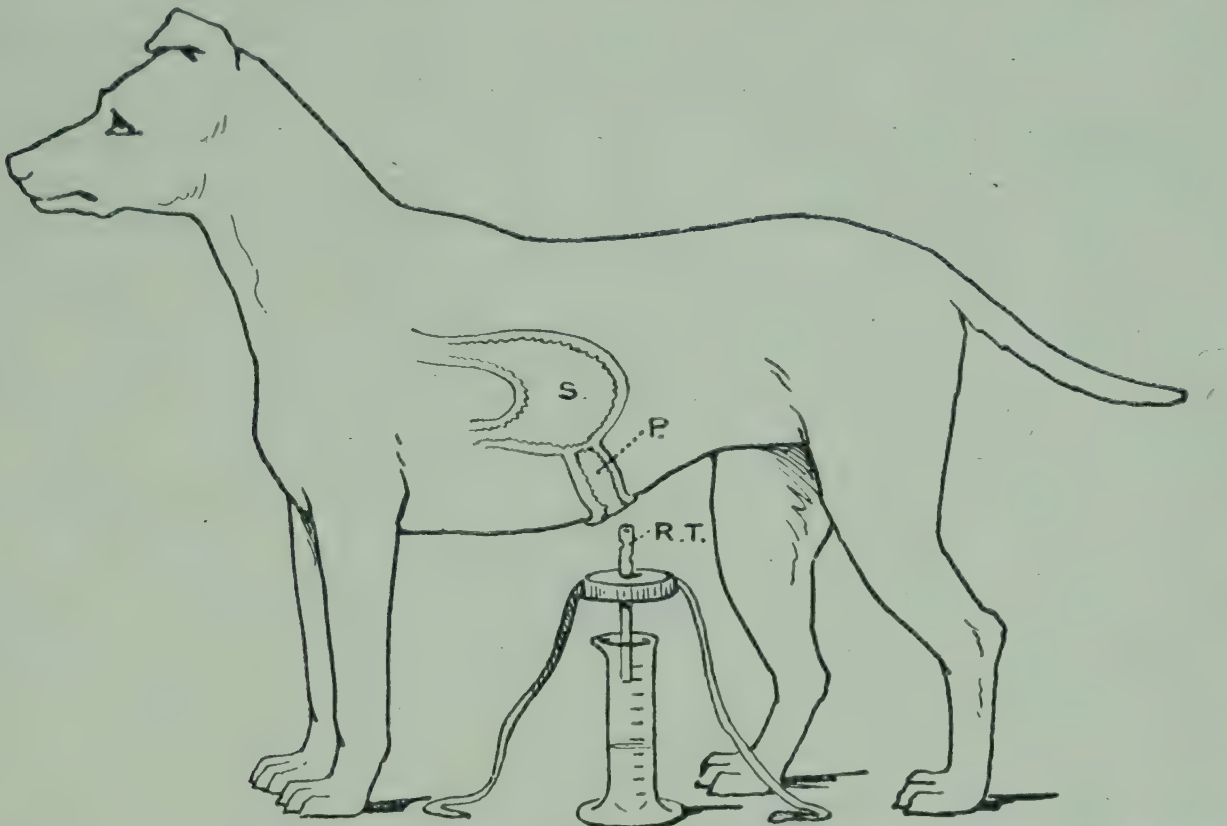


FIG. 560. Illustrating the Principle of the "Pavlov Pouch." S, body of stomach; P, the pouch; R.T., cannula made of indiarubber tubing, which is inserted into P, and held in position by tapes when juice is to be collected.

PSYCHICAL SECRETION OF GASTRIC JUICE (PAVLOV)

Period.	Quantity.
8 minutes	10 c.c.
4 ,,	10 ,,
4 ,,	10 ,,
10 ,,	10 ,,
10 ,,	10 ,,
8 ,,	10 ,,
8 ,,	10 ,,
19 ,,	10 ,,
9 ,,	3 ,,
SHAM FEEDING	
17 ,,	10 ,,
9 ,,	10 ,,
8 ,,	10 ,,

The efferent nerves in the reflex arc are the vagi; if they are divided the reflex disappears (this has also been shown in man); the afferent channels may be either the afferent nerves from the mouth or, in the conditioned reflex, any of the nerves of special sense, such as sight, which have, by being brought into use in association with the unconditioned reflex, acquired the properties of conditioned afferents. Or the action may be central in origin, as when the blood sugar has been lowered by insulin (7 units).

Chemical (Humoral) Secretion. The reflex secretion elicited by sham feeding soon declines after the feeding ceases. After a normal meal, the secretion of gastric juice continues for three to four hours, and much more juice is formed than by sham feeding alone.

It is thus evident that there are two phases of secretion of the gastric juice, the first a reflex phase occasioned by seeing or eating the food, the second due in some way to the presence of the swallowed food in the stomach. There has been some difference of opinion as to the manner in which this second phase is brought about. One experiment, however, settles an important point : a totally denervated pouch can be made, and transplanted into the mammary region ; if a dog so prepared is fed, the isolated pouch begins to secrete some considerable time after the animal has been fed, and the secretion continues for an hour or two.* This proves beyond all doubt that the second phase is not reflex, but is due to a blood-borne or humoral stimulus.

Several substances are known which will cause secretion of gastric juice, for instance, alcohol, soap, meat or liver extracts and peptone, when given by mouth, or histamine or pilocarpine, when given by intravenous injection. It might be supposed, therefore, that the secretion of the chemical phase was due to the absorption, into the blood, of some constituent of the food which had a secretagogue action. But it is not so simple as that : if a dog is furnished with a gastric fistula and also with a Pavlov pouch, and if meat is introduced through the fistula into the main stomach without the animal being aware of it, there will, of course, be no reflex secretion : but it will be found that there is no delayed secretion either. If the experiment be repeated, but with the difference that, together with the meat, some fresh gastric juice be also introduced into the main stomach, then, although the reflex secretion still remains in abeyance, the second phase secretion begins after the usual interval of time, and continues for a normal period. It is clear, therefore, that the humoral phase depends in some way on the presence, together in the stomach, of food and gastric juice. It must be the products of the action of the gastric juice on the food which act as secretagogues ; that this is so is proved by digesting the meat *in vitro* with gastric juice, and then giving the solution by stomach tube, when secretion begins almost at once.

The humoral secretion can be elicited by the contact of the exciting substances (*e.g.* liver extract or predigested meat) for a sufficient time with an adequate area of the mucosa of any part of the stomach, or of the small intestine This has been shown as follows : a dog was provided with a transplanted (*i.e.* a denervated) pouch, and the remainder of the stomach, with or without the pyloric part, was then isolated, denervated, and made into a large pouch : the lower œsophagus was connected directly to the duodenum. When the dog wa fed, or given liver extract orally, both stomach pouches secreted, *i.e.* the humora agent can function via the small intestine. When liver extract was introduce into the main pouch, the small transplant, as well as the main pouch, secreted But if the mucosa of the main pouch was anæsthetised with procaine, n secretion was obtained from either pouch when the main stomach was treate with liver extract, although on giving the extract by mouth, secretion followe in both, as usual.†

Nature of Humoral Excitant. The most obvious suggestion would be that th potent humoral agents are digestion products of proteins ; this would fit all but a few the facts. But the experiment just mentioned suggests that the active substance migh be a hormone, liberated into the blood from the gastric or intestinal mucosa when th chemical excitants formed by partial digestion of the food are brought in contact with it, the liberation of the hormone being inhibited if that mucosa had been previously treated with procaine. A claim to have identified a gastric hormone was, in fact, made by Edkins in 1906 [*]; he gave the name "gastrin" to the supposed hormone, and stated that it was only engendered from pyloric mucosa; but attempts to repeat his experimental results have been unsuccessful; so also were attempts to elicit gastric secretion when various excitants were placed in isolated pyloric pouches.[†] It seems established, nevertheless, that a histamine-free substance can be extracted from pyloric mucosa which, when injected intravenously, provokes secretion of gastric juice [‡]; the substance is a protein, rather resembling secretin, and it excites secretion of HCl, but not of pepsin.[§] It remains to be proved, however, whether this is a true hormone, *i.e.* whether that substance is actually taken up from the mucosa by the blood stream.

---

* Ivy and Farrell. *Amer. J. Physiol.*, 1925, **74**, 639.

† Gregory and Ivy. *Q. J. Exp. Physiol.*, 1941, **31**, 111.

It is unlikely that histamine is identical with the gastric hormone because the action of the chemical excitants, but not that of histamine, is inhibited by atropine. It is possible that histamine is liberated in the cells by the action, either of acetyl choline from vagus stimulation (abolished by atropine), or by the presence of the gastric excitant (or hormone).[‖]

Effect of Distension. The mechanical stimulus of distension causes secretion of acid gastric juice.[¶] The response must be caused by a hormone, since it still occurs when all nervous channels are eliminated. This has been shown by transplanting the pyloric part subcutaneously; when this is distended the body of the stomach secretes juice actively.[**]

Enterogastrone. The second phase of the gastric secretion is largely influenced by the character of the contents of the stomach. Thus after the ingestion of large quantities of fat, the secretion of gastric juice during the second phase is greatly delayed and also reduced in amount, in acidity, and in pepsin content. Emptying of the stomach is delayed, and there is considerable regurgitation of duodenal contents into the stomach, so that the volume of the gastric contents is subsequently increased. The inhibitory effect of the fat seems to be exerted from the duodenum and not from the stomach itself, as may be shown by the results of introducing fat into stomach or duodenum when these have been separated from one another. The effect is a humoral one, and the hormone concerned has been named *Enterogastrone*; it reduces volume and acidity of the juice resulting from a given stimulus: the pepsin content is also reduced when fat is administered, but this is a vagal effect.[††] Enterogastrone also reduces the motility of the stomach. After the preliminary delay, gastric secretion commences, often three or four hours late, and may then be quite brisk. This secretion is due to the action of the soaps which have been formed in the meantime by digestion of the fat by the regurgitated duodenal contents. A similar substance, *Urogastrone*, probably an excretory product of enterogastrone, has been extracted from urine. The main conclusions on the action of fat have been confirmed on man, *e.g.* when oil is introduced directly into the duodenum, or given before the meal, the inhibition of gastric secretion is greater than when it is swallowed with or after a meal.[‡‡]

The Normal Course of Secretion. We may say that the gastric secretion in response to a normal meal shows three phases: (1) a *reflex phase*, which begins within five minutes of the taking of the food and is determined by the reflex nervous mechanism described above; and (2) a *second phase*, of humoral nature, of later onset and longer duration, due to the presence of products of the digestion of the food in the stomach; (3) a *third phase*, small and variable, and due to the presence of digestion products in the duodenum. The normal gastric secretion is mainly due to the co-operation of the first two factors. The first and most important is the reflex secretion, determined through the vagus nerves by stimulation of the mucous membrane of the mouth, or by conditioned reflexes. The second factor, which provides for the continued secretion of gastric juice long after the reflex effects of feeding have disappeared, depends on the formation of an agent which acts as a chemical messenger to all parts of the stomach, being absorbed into the blood and thence exciting the activity of the gastric glands.

---

[*] Edkins. *J. Physiol.*, 1906, **34,** 133.

[†] Ivy and Whitlow. *Amer. J. Physiol.*, 1922, **60,** 578.

[‡] Komarov. *Rev. Canad. de Biol.*, 1942, **1,** 191, 377.

[§] Uvnäs. *Acta physiol. Scand.*, 1945, **9,** 269.

[‖] Babkin *Can. Med. Assoc. J.*, 1938, **38,** 421.

[¶] Lim, Ivy and McCarthy. *Q. J. Exp. Physiol.*, 1925, **15,** 13.

[**] Grossman, Robertson and Ivy. *Amer. J. Physiol.*, 1948, **153,** 1.

[††] Gray *et al.* *Amer. J. Physiol.*, 1937, **118,** 463; Grossman *et al.* *Amer. J. Physiol.* 1944, **141,** 281.

[‡‡] Roberts. *Quart. J. Med.*, 1931, **24,** 133.

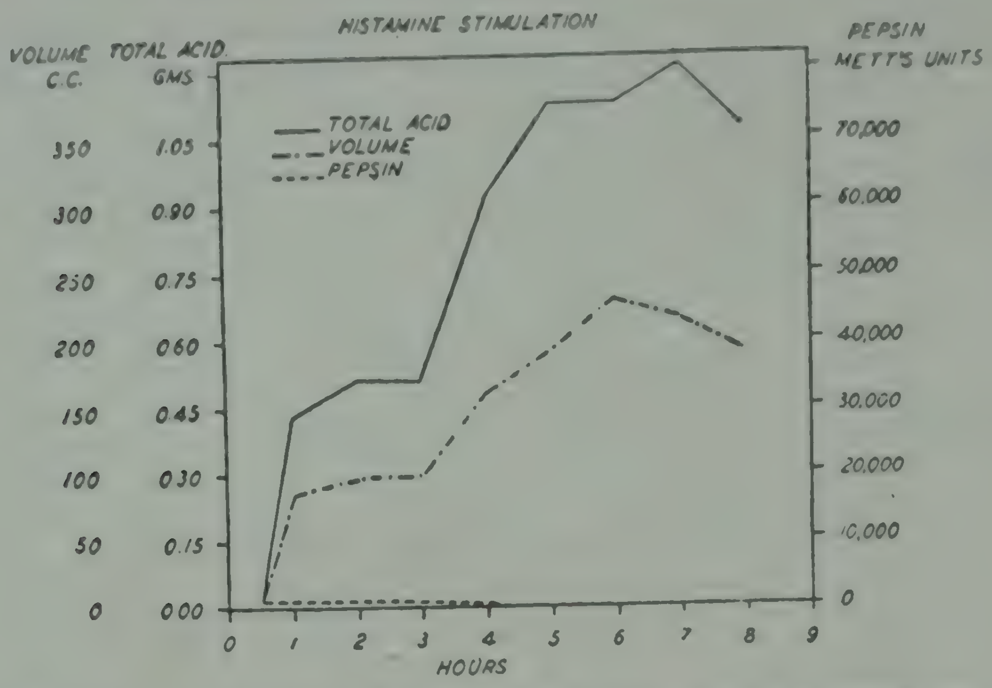


FIG. 561. Gastric juice resulting from hourly injections of histamine. The pepsin falls to zero, while acid and volume increase. (BOWIE and VINEBERG. *Q. J. Exp. Physiol.*, 1935, **25,** 247.)

It has long been known that the gastric juice is more acid the more rapidly it has been secreted. Pavlov believed that this was because there was a practically constant secretion of alkaline mucus, with which the gastric juice, of constant acid content, became mixed. This has been shown to be substantially correct by Gray and his co-workers.* They have also shown, by statistical examination of juices secreted at different rates that, in all probability, the parietal cells form a secretion containing 166 m.eq./l. of Cl, 7·4 of K, 158·6 of H-ion, while the non-parietal secretions contain 133·3 m.eq./l. of Cl, 33·0 of HCO_3, 154·5 of Na, 7·4 of K, and 3·7 of Ca. The composition of the juice can only vary between the limits set by these two components, and will be nearer to that of the parietal cells the faster the rate of formation.

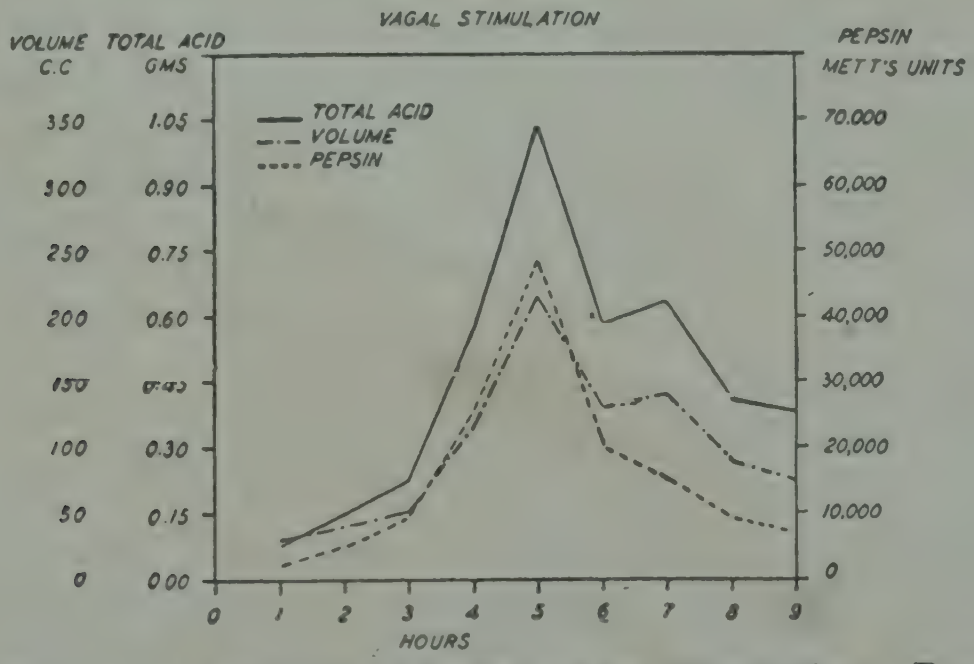


FIG. 562. Gastric juice from nine hours' vagus stimulation. Pepsin and acid are maintained. (BOWIE and VINEBERG.)

The acidity of gastric juice is constant, or nearly so, at about 0·135 N or 0·5 p.c. HCl, during the height of secretion. At the beginning and end, the juice is mixed with mucus, which dilutes and neutralises the acid.† Irritation of the mucosa by chemical or mechanical means increases the amount of mucus secreted.

A further important question is whether there is any alteration in the composition and amount of gastric juice with variations in the character of the food. So far as concerns the "appetite" juice, Pavlov showed that, whatever the previous diet, the juice always has the same characters. Since this "psychic" juice depends on appetite, it will be greater in amount the larger and the more welcome the meal. But in the case of the chemical phase, there is considerable variation in the nature of the juice. Whereas the volume of juice is greatest in amount after a meal of meat, the digestive power is greatest after one of bread. The juice secreted in this phase must vary according to the quantity of humoral agents produced, and therefore with the nature and amount of the substances produced in the preliminary digestion of the gastric contents, as well as with the initial composition of the food, and the composition and activity of the juice may vary with the intensity of stimulus. Fats reduce the gastric secretion and prolong the stay in the stomach. Water, tea or coffee taken with food augment gastric secretion, and various articles of food, such as vegetable juices, contain powerful chemical excitants (*e.g.* histamine in spinach).

---

* GRAY and BUCHER. *Amer. J. Physiol.*, 1941, **133,** 542.

† BOLTON and GOODHART. *J. Physiol.*, 1931, **73,** 115; HELMER, *Amer. J. Physiol.*, 1934, **110,** 28.

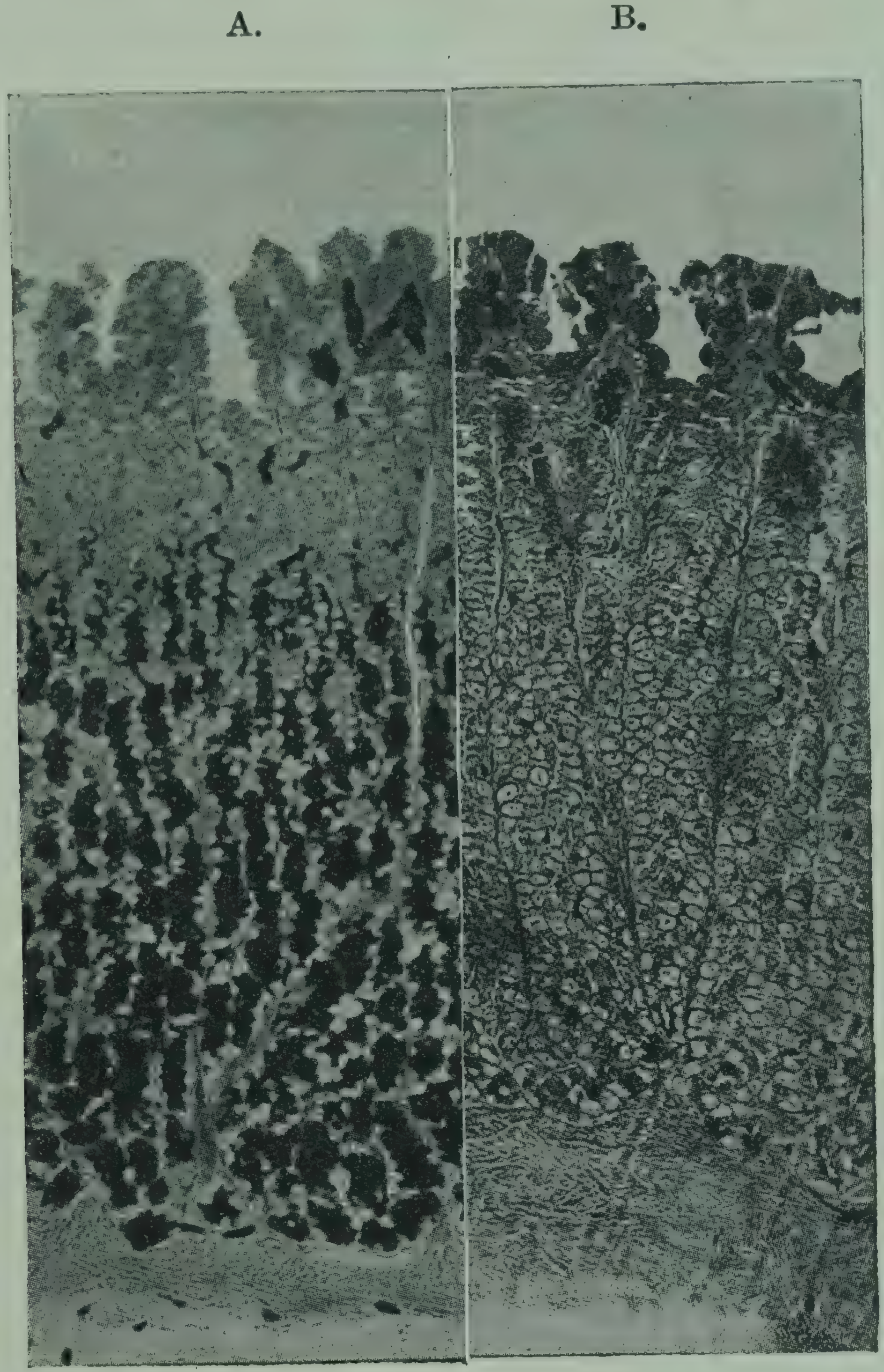


FIG. 563. A. Histological appearance of gastric mucosa after eight hours of histamine injection. The pepsinogen content is normal. B. Histological appearance of gastric mucosa after nine hours' vagal stimulation. The pepsinogen granules have nearly all vanished. (BOWIE and VINEBERG.)

According to Babkin * the vagus or reflex secretion is responsible for the control of the mucus, acid and pepsin, while the water and acid are mainly affected by the humoral control. These views have received strong confirmation from the work of Bowie and Vineberg.† They showed that when gastric secretion was excited for several hours by repeated injections of histamine, the pepsin content of the juice soon fell to zero, while the acid secretion continued (Fig. 561). Prolonged vagal stimulation, however, gave a juice rich in both acid and pepsin (Fig. 562), and vagus stimulation also caused the appearance of pepsin in the juice from which it had vanished after induction of secretion by histamine. Pepsinogen was found to be still abundant, as in the resting stomach, after prolonged secretion by histamine, whereas after prolonged vagal secretion the pepsinogen granules had almost entirely disappeared (Figs. 563 A and B). The juice secreted in response to pilocarpine is especially rich in pepsin, when compared with that provoked by histamine[*]: it also contains much mucin, which may be due to powerful contractions.[†]

---

* BABKIN. *Canad. Med. Assoc. Journ.*, 1931, **25**, 134.

† BOWIE and VINEBERG. *Q. J. Exp. Physiol.*, 1935, **25**, 247.

A possible collaboration of nervous and humoral factors has been suggested by Uvnas,[‡] who claims that vagal secretion is diminished if the pylorus is removed or cocainised: cross-circulation experiments show that during vagus stimulation a secretagogue, derived from the pylorus, enters the blood stream; the agent is presumably identical with that responsible for the gastric phase of secretion, and excites HCl secretion.

For Reference

Carlson, A. J. (1923). The Secretion of Gastric Juice in Health and Disease. *Physiol. Rev.*, **3**, 1.
Ivy (1941). The Mechanism of Gastric Secretion. *Surgery*, **10**, 861.
Pavlov (1910). The Work of the Digestive Glands.

4. THE MOVEMENTS OF THE STOMACH

These can be best studied by observation by means of the X-rays.[§] In order to make the shape of the stomach visible, the food—bread and milk—is mixed with a quantity of barium sulphate. The presence of this does not interfere with the processes of digestion, but renders the gastric contents opaque to the X-rays.

In the human stomach the term *fundus* is limited to that part of the stomach situated above the cardiac orifice (in the erect position). The *body* of the stomach is marked off from the pyloric part by the *incisura angularis* on the lesser curvature, represented in many animals by a strong " transverse band " (Fig. 564). The pyloric portion consists of the *pyloric vestibule* (or *antrum*) and the *pyloric canal*, the latter being a tubular portion with thick muscular walls about 3·0 cm. in length. A somewhat thicker ring of plain muscle surrounds the canal at its junction with the duodenum, and is called the *pyloric sphincter.*

When food has been swallowed (in the erect position) its weight is sufficient to overcome the resistance of the contracted gastric wall and some of it rapidly passes to the pyloric part. The remainder stays in the body of the stomach, which keeps constant pressure on its contents, tending to force them towards the pylorus. Contraction waves begin almost at once, each constriction starting near the middle of the stomach, and deepening as it slowly progresses towards the pylorus (Figs. 564, 565). These waves succeed one another, so that the pyloric part may present a series of constrictions. Their effect is to force towards the pylorus the food which has been mixed with gastric juice and detached from the surface of the mass in the corpus. If the pylorus is closed the food cannot escape, but it is doubtful whether the pyloric sphincter remains as persistently contracted as was formerly believed[||]; in fact it has been claimed that it is normally relaxed, closing only after a wave has passed over it from pyloric antrum to duodenum, this closure preventing reflux from the duodenum.[¶] The pyloric contractions last throughout the whole period of gastric digestion, and become more marked as it proceeds. By their action a thorough mixture of food and gastric juice results. Each wave as it travels along exhibits at intervals of a few seconds the rhythmic fluctuations in intensity common to most plain muscle contractions, and when a wave at a phase of increase reaches a point about one inch from the pyloric canal it results in part of the pyloric vestibule being nearly shut off from the rest of the stomach. The part thus confined then diminishes in size in every direction.[**]

---

[*] Bucher *et al.* *Amer. J. Physiol.*, 1941, **132,** 698.

[†] Hollander and Stein. *Amer. J. Physiol.*, 1943, **140,** 136.

[‡] Uvnäs. *Acta physiol. Scand.*, 1942, **4,** Suppl. xiii, 1–86.

[§] Cannon. *Amer. J. Physiol.*, 1898, **1,** 359.

[||] Wilson. *Can. Med. Assoc. Journ.*, 1937, **36,** 15.

[¶] Meschan and Quigley. *Amer. J. Physiol.*, 1938, **121,** 350.

[**] M'Crea, McSwiney *et al.* *Q. J. Exp. Phys.*, 1924, **14,** 379.

part of its contents being forced through the pyloric canal, while the remainder escapes back into the stomach. The waves recur at regular intervals of fifteen to twenty seconds, and three or four are present simultaneously. They continue without cessation until the stomach is empty—from one to four hours after the meal, according to its bulk and composition.

Although the co-ordination between the tonic contractions of the fundus and the rhythmic contractions of the pyloric part may be carried out by the local nervous system—Auerbach's plexus—situated between the layers of the muscular coat, it is probable that the advancing waves of contraction observed in the antrum are myogenic, *i.e.* directly originated in and determined by the muscle fibres themselves. Cannon has shown that these movements persist after complete division of Auerbach's plexus by from two to six circular incisions carried through the muscular coat of the stomach.

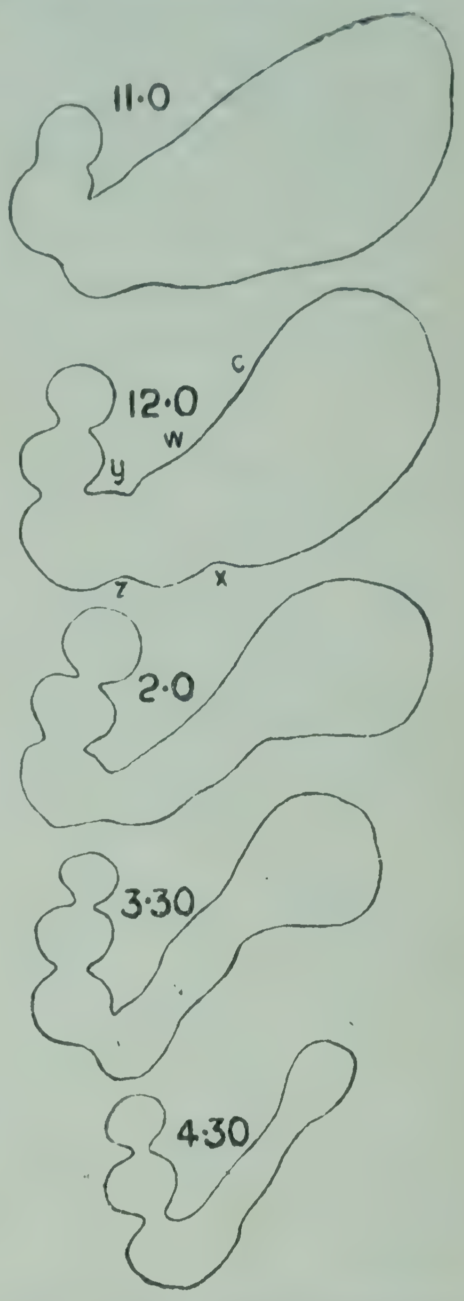


Fig. 564. Shadow Sketches of the Outlines of the Stomach of a Cat, immediately after a Meal (11.0) and at various intervals afterwards (12.0, 2.0, 3.30, 4.30). (W. B. Cannon.)

c. Situation of œsophageal opening.

yz. "Transverse band."

wx. Junction of body and pyloric portions.

The influence of the nervous system on stomach movements is undoubted, and emotional states or unpalatable food lead to suspension or reduction in the tone and movements of the stomach, a fact which should be borne in mind in interpreting X-ray findings after barium meals. Stimulation of the peripheral end of the vagus nerves may exercise varying effects on the gastric wall as well as on its sphincters. In the normal animal, stimulation of the peripheral end of the vagus as a rule causes strong contractions of the œsophagus as well as of the cardiac sphincter. After the administration of atropine, stimulation of the same nerve will occasion dilation of the cardiac sphincter. On both cardiac and pyloric portions of the stomach the vagus can exercise inhibitory as well as augmentor effects. So far as concerns the musculature of the fundus or body of the stomach, the most usual result is an inhibition during stimulation of the vagus, succeeded by an augmented tonus immediately the stimulus is removed. If the vagus be excited a number of times, the tonus of the muscular wall augments with each stimulus. On the pyloric portion, stimulation of the vagus also causes inhibition, followed by contraction and augmented peristaltic waves. McSwiney and Wadge * found that the effect of vagus stimulation on the stomach as a whole depends on its state of tonus. When the tonus is low, stimulation of the vagus causes contraction; when it is high, relaxation results. In the pylorus itself we may obtain from vagal stimulation either increased or diminished contraction.

Whether the splanchnic nerve has a direct influence on the movements of the stomach has been disputed. Most observers regard the splanchnic as having an influence on the stomach similar to its action on the intestine, and regard it as the chief inhibitory nerve to this organ.

Emptying of the Stomach. In the fasting stomach the pyloric sphincter is normally relaxed, and the pressures in the stomach and duodenal bulb are equal. When food is in the stomach, the pressure within it is generally slightly above that in the duodenal bulb.† As peristaltic waves pass along, the intragastric pressure first rises and then falls.

The mechanical factors concerned in the emptying of the stomach have been followed by studying the pressures in the pyloric antrum, within the sphincter, and in the duodenal bulb (first part of duodenum).

The principal factor in the emptying of the stomach is the passage of waves of contraction along the antrum.* During emptying, these waves pass uniformly along the antrum, sphincter and duodenal bulb,† and hence, when the pressure gradient is favourable, evacuation of some contents follows.‡ The pyloric sphincter plays a more subsidiary rôle, and is more important in preventing regurgitation of duodenal contents than in regulating the onward passage of contents from the pylorus.§ As emptying proceeds, the rate of evacuation is readily slowed down by influences, probably reflex, but also in part hormonal, which operate from the duodenum and upper intestine. The chief influences are those of large amounts of fluid, or the presence of fats, HCl, peptones,‖ sugars, or hypertonic solutions; all of these cause an inhibition of the tonus and movements in the antrum, sphincter and bulb. In emptying, the more fluid portions of the chyme are passed on first, and larger lumps retained for further digestion.

---

* McSwiney and Wadge. *J. Physiol.*, 1928, **65**, 350.

† Thomas. *Amer. J Physiol.*, 1938, **123**, 201.

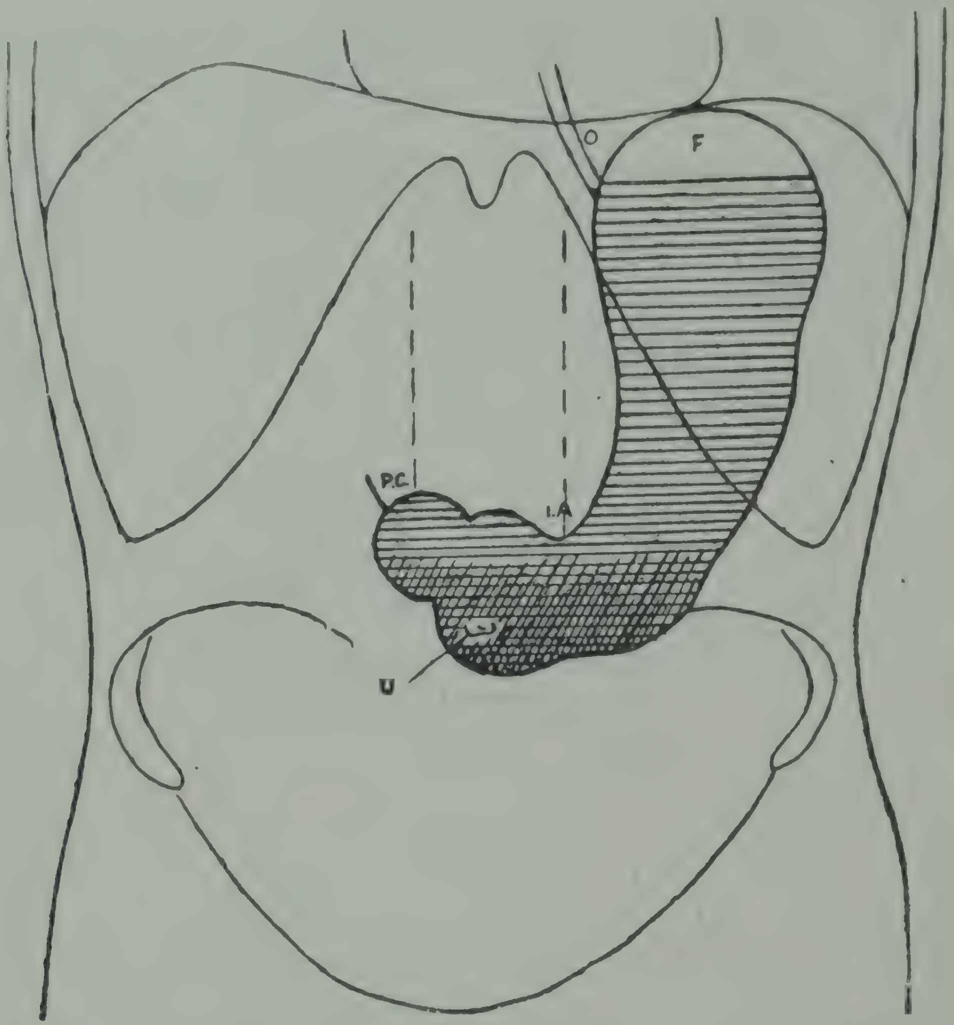


FIG. 565. Sketch of Human Stomach, in Erect Position, shortly after a Bismuth Meal. (HURST. *Quart. J. Med.*, 1910, **3,** 374.)

F. Fundus. U. Umbilicus. IA. Incisura angularis.
PC. Pyloric canal. O. Œsophagus.

The normal emptying of the stomach is also largely dependent on the integrity of the connection of this viscus with the central nervous system. Hence if, in a dog, both vagus nerves be divided below the point at which they give off their branches to the lungs and heart, a large amount of food may remain in the stomach in an undigested condition. The secretion of gastric juice is deficient, and the opening of the pylorus is not easily carried out.

Quantitative study of the emptying of the human stomach has been made by the use of "serial test meals."¶ To obtain basal data an indifferent substance such as pectin solution was used: 750 c.c. of the solution, dyed with phenol red and of fixed pH and osmotic pressure were given daily. Each day the whole contents were withdrawn, the time of withdrawal being varied each day, and the results were synthesised into one "serial meal" which gives the normal sequence of emptying. The volume of gastric contents passing the pylorus between any two stages is got from the amount of dye lost by the contents between the stages. In this way Fig. 566 was obtained. Fig. 567 derived from this, shows that the rate of emptying stands in linear relationship to the volume of the gastric contents. A further calculation can be made to give the volume of secretion added to the contents.

---

* THOMAS. *Rev. Gastroenterol.*, 1935, **2,** 32.

† MESCHAN and QUIGLEY. *Am. J. Physiol.*, 1938, **121,** 350.

‡ WERLE *et al.* *Am. J. Physiol.*, 1941, **131,** 606.

§ QUIGLEY. *Arch. Surg.*, 1942, **44,** 414.

‖ VAN LIERE, SLEETH and NORTHUP. *Amer. J. Physiol.*, 1937, **119,** 480.

¶ HUNT and SPURRELL. Communic. to Physiol. Soc., February 19th, 1949.

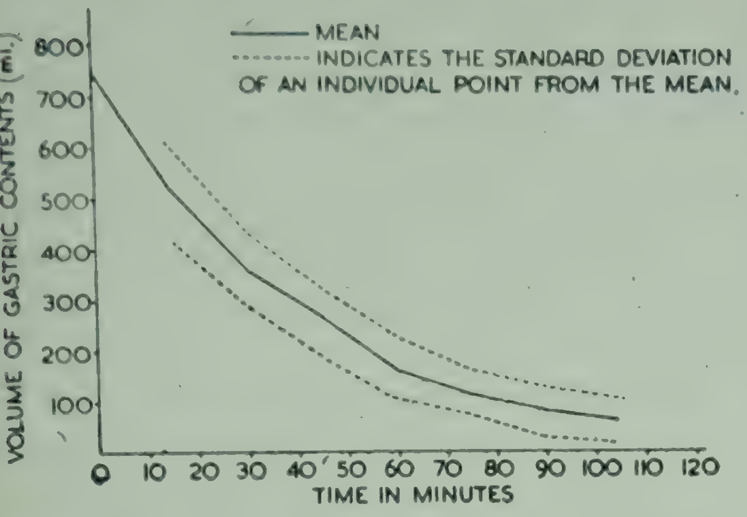


FIG. 566. Volume of gastric contents plotted against time, as determined by 185 serial test meals on 19 healthy students. (Courtesy of J. N. HUNT and W. R. SPURRELL.)

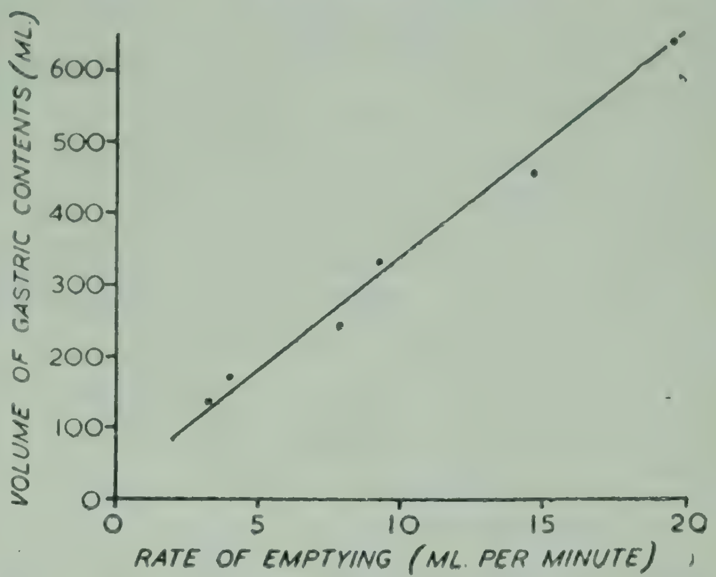


FIG. 567. Rate of emptying of gastric contents plotted against volume of gastric contents. (Courtesy of J. N. HUNT and W. R. SPURRELL.)

Liquids leave the stomach earlier than solid material. If fluid alone be swallowed, as when water is taken to quench thirst, it may arrive in the duodenum within a minute or two after being swallowed. An increase of duodenal pressure causes increased tone in the pyloric sphincter, and so retards emptying of the stomach, while a lowering of duodenal pressure has the reverse effect.* After larger meals the emptying time is greater in most subjects than after small meals, *e.g.* doubling a small test meal may increase the time by 17 per cent., and trebling it may increase it by 38 per cent.†

Meals of different foodstuffs remain in the stomach for different periods. Carbohydrates leave early, whereas with fats there is longer retention, the delay being proportional to the quantity of fat,‡ and produced when fat enters the upper intestine; it is humoral in origin, and due to liberation of enterogastrone (p. 867). It would follow from this, and indeed, it has been proved— that when fat is taken with a meal, the stomach empties so slowly that the duodenal contents never become more than slightly acid, whereas in absence of fat the pH falls considerably during emptying of the stomach.

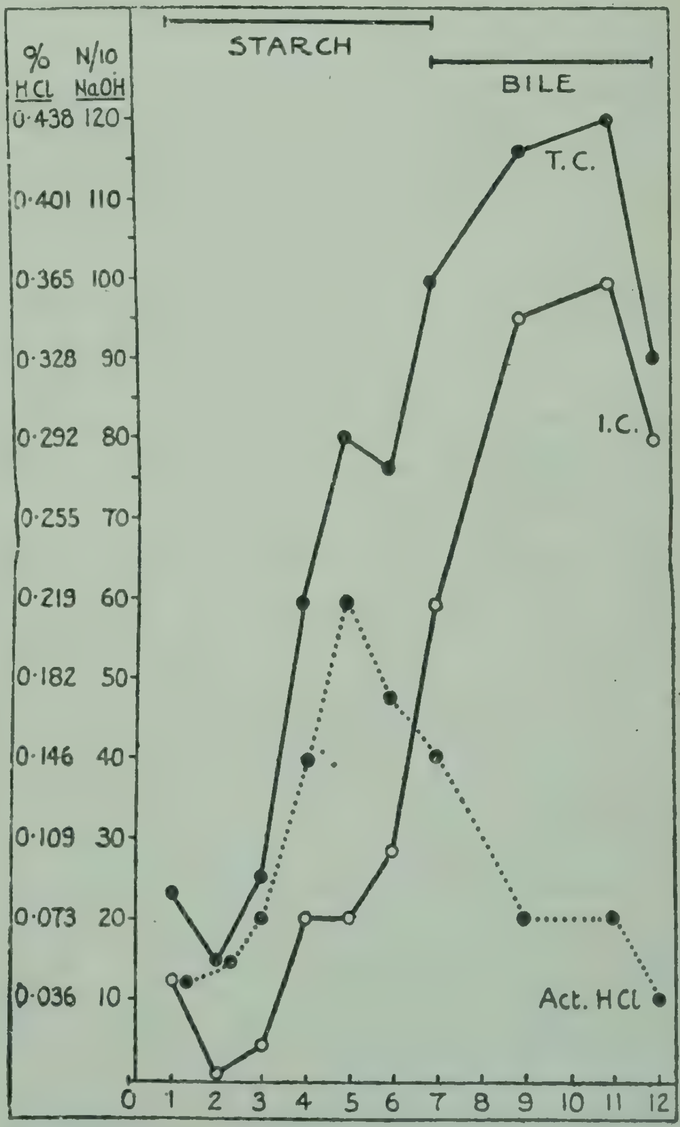


FIG. 568. Analysis of Gastric Contents at 15 min. intervals after a Test Meal of Gruel. T.C., total chlorides; I.C., inorganic chlorides. Act.HCl., "active" hydrochloric acid. (BOLTON.)

Towards the end of digestion a regurgitation of duodenal contents into the stomach will occur if the pressure is higher in the duodenum than in the antrum. The normal pressure in the duodenum is about 15 cm. water, and the pyloric intragastric pressure only 10 cm. during relaxation, but 20 to 30 cm. during an intense wave.

---

* SPURRELL. *J. Physiol.*, 1935, **84**, 4.P.

† CRIDER and THOMAS. *Amer. J. Physiol.*, 1938, **123**, **44**.

‡ MCSWINEY and SPURRELL. *J. Physiol.*, 1935, **84**, 41.

The contractions of the duodenum itself play an important part in regulating the emptying of, and regurgitation into, the stomach. Forward contractions empty the duodenal cap and facilitate exit from the stomach. Anti-peristalsis also occurs and, by filling the duodenal cap, gives an opportunity for regurgitation (Bolton).

The exact parts played in this mechanism by the local nerve net and by the central nervous system respectively have not yet been thoroughly made out, though there is no doubt that some movements may proceed independently of any connection with the central nervous system.

When the stomach is empty the pylorus is kept relaxed, admitting of the free entry of duodenal contents (pancreatic juice and bile) into the stomach. But this regurgitation, especially of pancreatic juice, seems to be a frequent phenomenon even during digestion, and particularly when the meal contains fat, and to be a means by which the acidity of the gastric contents is prevented from rising above an optimum of 0·15 to 0·2 per cent. HCl.

Fractional examination of the gastric contents shows that, normally, the acidity reaches a maximum and then declines ; after the decline has begun, bile is present in the fractions, showing that regurgitation has occurred (Fig. 568). If regurgitation is prevented in any way, the contents do not show the declining acidity, and no bile is found (Bolton).

The " *pangs of hunger* " are associated with, and are probably due to, prolonged rhythmic contractions of the stomach wall, which come on as soon as the stomach is empty, and increase in force and frequency ; each contraction lasts about 20 seconds, and is accompanied by salivation and augmentation of the knee-jerks.* The hunger contractions are inhibited by the introduction of food into the stomach, by the act of swallowing and also by intravenous injection of glucose. Administration of insulin, which lowers the blood sugar, augments the contractions and causes hunger pangs.†

For Reference

McSwiney (1931). Innervation of the Stomach. *Physiol. Rev.*, **11,** 478.

Vomiting. Rejection of the stomach contents may occur as a result of over-distension, especially of the pylorus, of the presence of irritating contents, or reflexly from various organs. It is generally preceded by a feeling of nausea, associated with excessive salivation, and other signs of central autonomic stimulation, *e.g.* pupil dilatation, sweating, pallor, etc. In the act of vomiting itself, the first event is a deep inspiration. The glottis is then closed and the nasal chamber often shut off by raising of the soft palate, and this is followed by a strong contraction of the diaphragm and of the abdominal muscles. At the same time the œsophagus, the body of the stomach and the cardiac orifice are relaxed, and the tone and movements of the upper intestine first inhibited, and then augmented.‡ By means of X-rays it is seen that at this time a strong contraction occurs at the incisura angularis, dividing the stomach into two separate portions. The dilated body of the stomach is pressed between the abdominal muscles and the diaphragm, so that its contents are expelled. Contraction of the body of the stomach may be slight or absent and the main factor in the expulsion is the contraction of the abdominal muscles and diaphragm. In fact, vomiting may be excited in an animal in which the stomach has been replaced by a bladder. Finally, the œsophagus is emptied, partly by the positive pressure in the chest as a result of the action of the abdominal muscles, partly by contraction.

---

* Cannon and Washburn. *Amer. J. Physiol.*, 1912, **29,** 441 ; Carlson, " The Control of Hunger in Health and Disease," 1916. (Chicago Univ. Press.)

† Carlson and Quigley. *Amer. J. Physiol.*, 1931, **97,** 107.

‡ Gregory. *J. Physiol.*, 1946, **105,** 58.

Nervous Mechanism of Vomiting. Normally the action of vomiting is reflex. It can be excited by tickling the back of the throat, when the afferent nerves are the trigeminal and the glossopharyngeal, or by irritation of the stomach, through the afferent fibres of the vagus. But it may be excited reflexly from almost any of the abdominal viscera, *e.g.* uterus, kidney, intestines, etc., or through the labyrinth or eyes, as is supposed to be the case in the vomiting of sea-sickness, and it is a marked symptom in many cases of disease of the cerebrum or cerebellum. The efferent impulses are carried by the vagi to the stomach, by the phrenics to the diaphragm, and by the various spinal nerves to the abdominal muscles. There are also inhibitory impulses descending the vagi to the œsophagus and cardiac sphincter, and inhibitory, followed by excitatory, impulses to the intestines.* The reflex act depends on the integrity of the medulla (" vomiting centre "). Drugs may produce vomiting either reflexly, by irritating the stomach, *e.g.* mustard and water, or by direct action on the medullary centres, *e.g.* apomorphine, etc.

For Reference

Hatcher, R. A. (1924). The Mechanism of Vomiting. *Physiol. Rev.*, **4**, 479.

5. INTESTINAL DIGESTION

The contents of the stomach are passed at intervals into the first part of the duodenum. Here they meet the secretions of the pancreas, the liver, the tubular glands of the intestine, and the secretion of Brunner's glands. The acidity is partly neutralised in the duodenum, the contents of which have a fasting pH of about 6·5, which may fall to 3·5 during gastric emptying ; further reduction of acidity occurs lower down in the intestine, the pH in the ileum being about 6·0 to 6·5, though alkalinity occurs rarely if at all.

The secretion of these various juices is simultaneously aroused by the entry of the acid chyme into the duodenum. Although they co-operate in their action on the foodstuffs, it will be convenient to deal separately with each.

Duodenal Contents. Samples of the duodenal contents may be obtained in man by passing an Einhorn's duodenal tube, a 30-in. long narrow rubber tube with a small lead plummet embedded at the end. This is swallowed, and when the plummet reaches the stomach the patient lies on the left side until it reaches the duodenum, where its position can be verified by X-ray examination. By suction with a syringe the contents, frothy, and yellow, are withdrawn. If 20 c.c. of 50 per cent. $MgSO_4$ is injected down the tube a copious flow of bile follows.

The Pancreatic Juice

The Secretion of Pancreatic Juice. The normal secretion of pancreatic juice can be studied on an animal with a pancreatic fistula.

In order to establish a pancreatic fistula in a dog, a small piece of the duodenal wall is exsected, having the papilla of the lower (and larger) duct opening in the middle of its mucous surface. The integrity of the gut is restored, and the exsected piece is brought to the surface and stitched in the middle of the abdominal wound. The greater part of the pancreatic secretion will escape by the fistula, and can be collected. An alternative and more modern method is to tie a glass cannula into the pancreatic duct, and to it attach a rubber tube which is led out through the abdominal wall. The continual loss of pancreatic juice may seriously affect the animal's health. This may be obviated to a certain extent by keeping the animal on a milk diet with the addition of sodium bicarbonate to replace the loss of base by the juice.

Clinical observations of human pancreatic fistulæ have confirmed the experiments made on dogs.

In the fasting condition there is for long periods (one to two hours) no expulsion of juice from the duct. Probably, however, there is occasional slow secretion of juice, coincident with a phase of spontaneous periodic activity of the alimentary canal (Babkin), so that the escape of a few drops may be observed at long intervals. If a meal be administered, a flow of juice begins in one to two minutes. This first secretion is reflex. From this time there is a steady, slow rise in the rate of secretion, which lasts for two to three hours, and then gradually diminishes. The greatest increase in flow is observed at the time when the first portions of digested food escape from the stomach into the duodenum. The secretion must therefore be determined in some way by this entry into the duodenum. By experiments on dogs possessing a gastric as well as a pancreatic fistula, it has been shown that the introduction of acid, *e.g.* 0·4 per cent. HCl, into the stomach evokes, as soon as it passes into the duodenum, a rapid flow of pancreatic juice. A smaller effect is produced by the passage of oil from the stomach into the duodenum. The introduction of alkalies is without influence. Dilute acids are also effective exciters of secretion when introduced directly into the duodenum or small intestine. The secretion produced diminishes as the loop employed is nearer to the cæcum, and as a rule the injection of dilute acid into the lower foot or eighteen inches of ileum is without effect. This response to acid is not due to a reflex, but to a humoral (chemical) mechanism.

---

* Gregory. *J. Physiol.*, 1947, **106**, 95.

Nervous Secretion. The pancreas receives fibres from the vagi and from the splanchnic nerves. Pavlov noticed that if the vagus on one side were cut in the neck and left for four days in order to allow the cardio-inhibitory fibres to degenerate, prolonged stimulation of the peripheral end of the nerve evoked a slow flow of pancreatic juice. He obtained the same results by stimulating this nerve below the heart, in animals with the cervical spinal cord divided. Pancreatic secretion resulted from vagus stimulation when every care was taken to prevent acid stomach contents being expelled into the duodenum by contraction of the stomach wall ; *e.g.* when the pylorus was ligatured and the stomach filled with an alkaline solution. There can be no doubt that the vagus is a secretory nerve to the pancreas, and that the secretion of pancreatic juice which occurs within a minute or two of the taking of food is a reflex. The juice obtained by vagus stimulation is glairy, very concentrated in solids, and rich in enzymes. When secretion has been established by prolonged and repeated stimulation of one vagus, a stimulation of the other vagus at first checks it. This is due to constriction of the ducts.* A slight secretion, inhibited by atropine and potentiated by eserine, also occurs on stimulation of the splanchnic nerves, but not on injection of adrenaline. The splanchnic therefore contains some cholinergic fibres.†

Humoral Secretion. The injection of acid into a loop of small intestine is followed by secretion of pancreatic juice. Bayliss and Starling ‡ found that a secretion of juice was produced even when the acid was introduced into a loop of the small intestine entirely freed from any possible nervous connections with the rest of the body. It was evident, therefore, that the stimulus from the intestine to the pancreas, which causes the secretion of the latter, must be carried, not by the nervous system, but by the blood stream. The injection of acid into the portal vein was without effect, but on boiling some scrapings of the intestinal mucous membrane with dilute hydrochloric acid, neutralising and filtering, and injecting the filtrate a copious flow of pancreatic juice was produced. This classical experiment provided the first known instance of a chemical excitant, and the name " hormone " was given to this class of substances.

Secretin. This chemical messenger, or hormone, is called " secretin." That it is present in the mucous membrane is shown by the fact that secretin can be extracted also by water, solutions of soap, or dilute alkali, or by the prolonged action of alcohol upon mucous membrane which has been killed by heat. Neutral solutions, however, extract less secretin than do acid solutions.

---

* Anrep. *J. Physiol.*, 1916, **50**, 421.

† Babkin *et al.* *Q. J. Exp. Physiol.*, 1939, **29**, 217.

‡ Bayliss and Starling. *J. Physiol.*, 1902, **28**, 325.

This may be because some of the secretin is held in an adsorbed or otherwise "bound" form, from which it is freed by acids.*

Highly active preparations of secretin may be prepared in various ways. One † is by extraction of the duodenal mucosa with absolute alcohol, precipitation by dilute acid, and resolution of the precipitate in acid alcohol and precipitation by acetone. Various other methods have been used and crystalline preparations ‡ § may be obtained. Greengard and Ivy's crystalline secretin is obtained as a picrolonate of empirical composition C_3H_5ON; 0·014 mg. of the free secretin from this = 1 Ivy-dog-unit, *i.e.* it gives 1 c.c. of juice. It is a readily soluble substance resembling a protein or high polypeptide, is rapidly destroyed by pepsin and trypsin, and diffuses slowly through animal membranes. Though stable in acid solutions, it is very rapidly destroyed in alkaline or neutral solutions, especially under the influence of bacteria.

In this secreting mechanism we have an example of a correlation effected by chemical means, between the activities of two different portions of the body. The acid chyme enters the first part of the duodenum. Immediately a certain amount of secretin is extracted by the acid, is carried by the blood stream to the cells of the pancreas and excites there the secretion of strongly alkaline pancreatic juice. As soon as sufficient juice has been secreted to nearly neutralise the acid chyme, the formation of secretin, and therefore the further secretion of alkali, comes to an end. If the stomach still contains food, however, the process is renewed, the pylorus relaxes and allows the entrance of a further portion of acid chyme. Thus the whole chain of processes goes on until the stomach is empty and all its contents have passed into the intestine. One of the actions of the powerful secretin preparations made by Mellanby is that it causes contractions of the intestinal muscle; it also causes some bile secretion.

For Reference

STILL (1931). Secretin. *Physiol. Rev.*, **11**, 328.

"***Pancreozymin.***" Pure secretin produces a secretion of water and salts, but does not excite secretion of enzymes, though crude secretin preparations do so. Further, when a meal is given to a dog with a transplanted pancreas, the enzyme content, as well as the amount of the juice secreted, increase.||

The explanation appears to have been provided by the discovery that, as usually prepared, secretin extracts contain a substance which does excite the secretion of enzymes, but not of water and salts. The substance is called pancreozymin, and it and secretin can be prepared free each of the other. The two substances are extracted together by alcohol: the alcohol is distilled off, the secretin removed by bile-salts, and the pancreozymin precipitated by sodium chloride.¶ Pancreozymin appears to have similar chemical and physical properties to secretin. It accompanies secretin in the epithelium of the intestine. Like vagus stimulation, pancreozymin causes a discharge of zymogen granules from the cells of the pancreatic acini, which secretin alone does not. But pancreozymin acts on the atropinised pancreas, while vagus stimulation does not. Its action is thus probably a direct one on the cells.**

The secretin mechanism provides alkali for the neutralisation of the acid stomach contents. But the pancreozymin and the nervous mechanism provide the enzyme content of the juice, so that, even when for any reason the gastric contents are not acid, pancreatic digestion is still possible.††

In the absence of HCl from the gastric contents, it is doubtful what is the responsible agent in starting the secretin mechanism. The presence of bile or

---

* SCOTT and STILL. *Amer. J. Physiol.*, 1935, **112**, 511.

† J. MELLANBY. *Proc. R. S.*, 1932, **111B**, 429.

‡ HAMMARSTEN *et al.* *Acta. Med. Skand.*, 1928, **68**, 1.

§ GREENGARD and IVY. *Amer. J. Physiol.*, 1938, **124**, 427.

|| FARRELL *v.* IVY. *Amer. J. Physiol.*, 1926, **78**, 325.

¶ HARPER and RAPER. *J. Physiol.*, 1943, **102**, 115.

** HARPER and MACKAY. *J. Physiol.*, 1948, **107**, 89.

†† J. MELLANBY. *J. Physiol.*, 1925, **60**, 85; *Proc. Roy. Soc.*, 1932, **111B**, 429.

bile salts in the intestine does not augment the secretion of pancreatic juice, and may even reduce the response to acid, peptone or soap.*

The amount and activity of pancreatic juice obtained after a meal varies with the nature of the latter. The differences seem largely determined by the amount of acid passed on to the duodenum, and by the presence of specific excitants, such as fat and soap, which raise the enzyme content of the juice.

Electrical Changes. Stimulation of the pancreas by secretin injection causes electrical changes in which the duct of the pancreas becomes positive to the surface of the gland.†

THE STRUCTURAL CHANGES IN THE PANCREAS ACCOMPANYING SECRETION. A normal resting gland on section is seen to consist of numerous secreting alveoli, the cells of which present two well-marked zones: a narrow peripheral zone in which the nucleus is embedded, and which is strongly basophile; and a central zone which is turned towards the lumen, occupying two-thirds or three-quarters of the cell, and is closely packed with highly refractive granules, strongly oxyphile, and containing the zymogens of the pancreatic juice (Fig. 569 A). After the induction of secretion by stimulation of the vagus, the lumen of each alveolus is enlarged, the cells are shrunken, and the granules are found to lie only along the border of the cell turned towards the lumen, the rest of the cell, which is much reduced in size, being made up of the basophile cytoplasm (Fig. 569 B). After secretion has been caused by secretin injection, the changes are only slight.

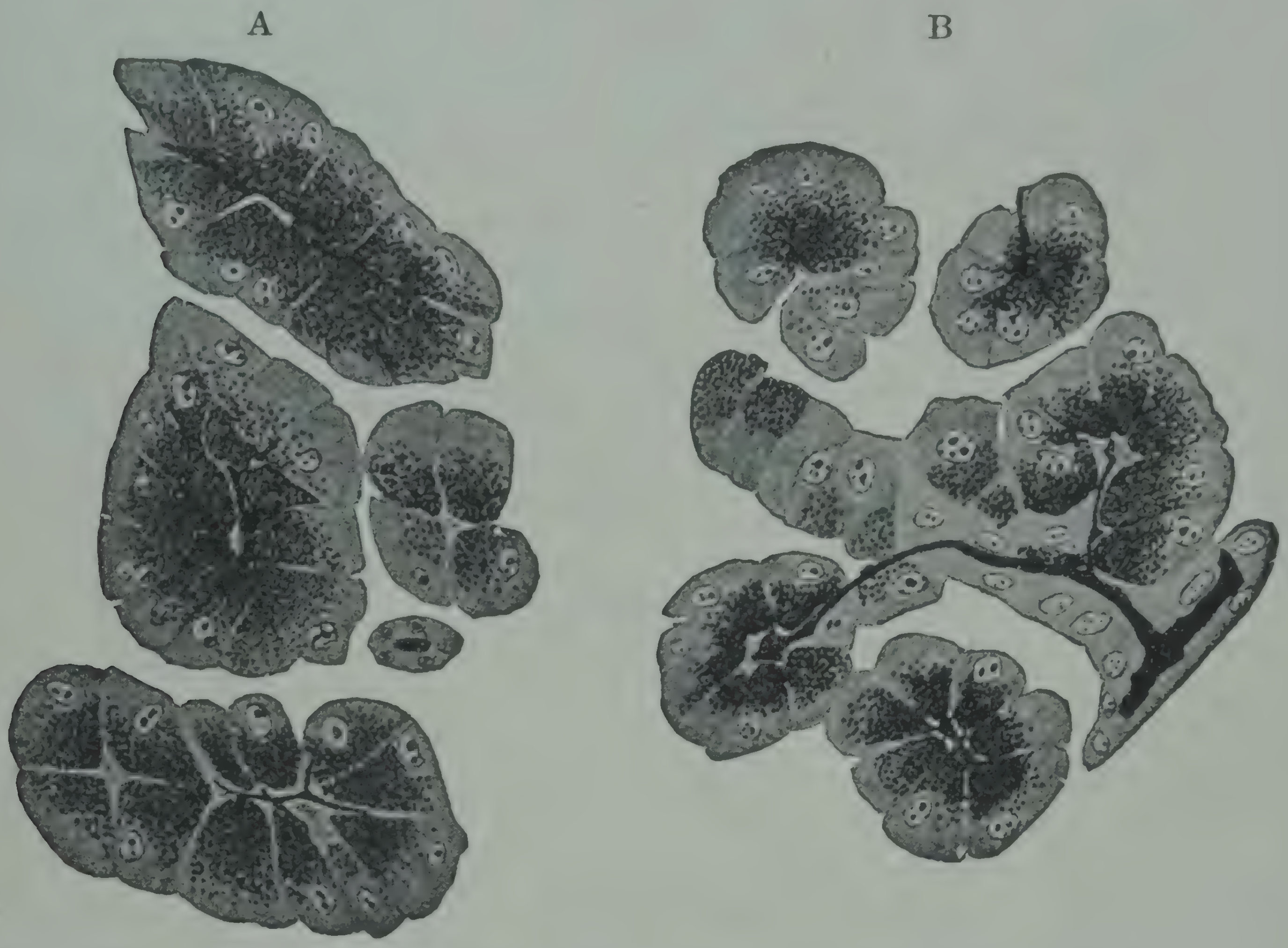


FIG. 569. Alveoli of Dog's Pancreas. (BABKIN, RUBASCHKIN and SAVITSCH, *Arch. Mikr. Anat.* 1909, **74,** 68.)

A. Resting. B. After moderate secretion with discharge of granules.

Properties of Pancreatic Juice. Pancreatic juice can be obtained by the injection of secretin or pilocarpine intravenously, or by stimulation of the vagus. The average composition varies according to the means used to cause secretion, as is shown in the table on p. 879.

"Secretin" juice is a clear or slightly opalescent fluid, alkaline ($p\text{H} = 7\text{–}8$) from the presence of sodium bicarbonate, its titratable alkalinity varying between $\frac{N}{10}$ and $\frac{N}{7}$ alkali, and therefore about equivalent to gastric juice in normality. The proteins include the enzymes and serum proteins. The juice tends to become poor in enzymes, but not in alkali, as secretion proceeds. The concentrated juice obtained by injection of pilocarpine, or by vagus stimulation, which

---

* THOMAS and CRIDER. *Amer. J. Physiol.*, 1943, **138,** 548.

† ANREP and DALY. *J. Physiol.*, 1921, **55,** ii, Proc.

	A	B		C
Alkalinity:		(a)	(b)	
Number of c.c. $\frac{N}{10}$ NaOH equal to 10 c.c. juice	12·7	12·4	9·0	5·5
i.e. in terms of Na in 100 c.c.	0·29	0·28	0·20	0·12
Total solids in 100 c.c.	1·6 1·56	2·25	1·5	6·38 6·40
Total proteins in 100 c.c.	0·5	—	—	4·8
Ash in 100 c.c.	1·00 0·92	1·00	1·00	1·3
Chlorides in 100 c.c.	0·28 0·30	—	—	0·27
Total nitrogen	—	—	—	0·74

A. Secretin juice from three dogs. Sp. gr. 1014.
B. Secretin juice, specimen collected at beginning (*a*), and at end (*b*).
C. Pilocarpine juice.

may contain as much as 6 per cent. total solids, is always considerably less alkaline and richer in proteins and enzymes than the more dilute juice got by injection of secretin.

The most interesting and important constituents of the juice are its enzymes. The juice on arrival in the intestine has an effect on proteins, fats and carbohydrates, due to the presence of distinct groups of enzymes, viz. trypsin, lipase and amylase.

The pancreas tissue contains the same enzymes as the juice, and in the same relative amounts.

The Activation of Pancreatic Juice. Freshly secreted pancreatic juice, collected without contact with the intestinal mucosa, has no digestive action on egg-white, though able to digest peptone, histone or protamines. Succus entericus very rapidly activates it, owing to the presence in it of *enterokinase.** Enterokinase acts like an enzyme † on the inactive enzyme (trypsinogen), splitting off a polypeptide and releasing the activated enzyme trypsin. Trypsinogen has been prepared in pure crystalline form,‡ and is rapidly activated by enterokinase at about *p*H 5·5.

Most of the controversy as to whether enterokinase is an enzyme appears to be due to the use of crude preparations.§ Pure crystalline trypsinogen is autocatalytically activated by the minutest trace of trypsin. Crude preparations, however, contain an inhibitor of trypsin which checks this reaction, so that normally only the kinase can start it, and this is the state of affairs in the alimentary tract, where activation occurs in a few seconds.

If pancreatic juice be allowed to stand, it gradually acquires activity. If, however, sodium fluoride be added, the juice remains permanently inactive. The spontaneous activation of the juice may be hastened by neutralisation. The most potent means next to enterokinase is the addition of calcium salts. According to J. Mellanby and Woolley ‖ the calcium acts simply by neutralising the juice and thus allowing minute traces of kinase already present in the juice to exert their effect; or it might act by diverting the trypsin inhibitor.

The Action of Pancreatic Juice on Proteins. Activation is effected directly the juice enters the gut, and it thus acquires a proteolytic activity superior to that of any other digestive juice.

The enzymes of the juice may be separated from one another, and so studied separately, by appropriate adsorption treatment as outlined in Chapter IX.

---

* Schepowalnikov. *Jahresb. Fortschr. Tierchem.*, 1900, **29**, 378.

† Vernon. *J. Physiol.*, 1901, **27**, 269; 1913, **47**, 325. Kunitz. *J. Gen. Physiol.*, 1939, **22**, 429.

‡ Kunitz and Northrop. *J. Gen. Physiol.*, 1934, **17**, 591.

§ Northrop. *Science*, 1937, **86**, 479.

‖ Mellanby and Woolley. *J. Physiol.*, 1913, **46**, 159.

Inactive pancreatic juice contains two zymogens, viz. chymotrypsinogen and trypsinogen.* Chymotrypsin clots milk, and hydrolyses casein and gelatine, but differently from trypsin; it does not hydrolyse polypeptides in general. Enterokinase converts trypsinogen to trypsin, but does not act on pure chymotrypsinogen. This is, however, acted upon by trypsin and converted into chymotrypsin. Hence, enterokinase added to a mixture of the two zymogens results in activation of both.

Trypsin acts on proteins generally, converting them to polypeptides.

It is now known that "trypsin" is a mixture of several proteolytic enzymes, *e.g.* trypsin, aminopolypeptidase, carboxypolypeptidase, protaminease, and no doubt others. It also contains a substance which inhibits trypsin. These enzymes act in different ways, or on different substrates. Thus carboxypolypeptidase does not act on proteins, but only on some of their hydrolysis products, from which it splits off a terminal amino acid which in the intact substrate had a free carboxyl group. The free amino group is not necessary and hence carboxypolypeptidase will act on peptic digests which have been treated with formaldehyde to remove free NH_2 groups.† Aminopolypeptidase only acts on polypeptides with free α-amino groups, splitting off the amino-acids.

The different constituents of the protein molecule show a varying resistance to the action of trypsin. Denaturation of the protein may be the earliest stage. The greater part of the molecule is rapidly broken down and yields proteoses, peptones, polypeptides, and amino-acids; and the same change is undergone by the proteoses and peptones resulting from the gastric digestion of proteins. Within a few minutes, therefore, after the chyme has reached the small intestine, a certain amount of amino-acids will have been formed. Trypsin thus differs fundamentally in its action from pepsin, which never yields amino-acids. Some of the products present a resistance to disintegration. This part gives the biuret reaction. Even when the trypsin has acted *in vitro* for weeks, and the biuret reaction has entirely disappeared, the mixture will be found to contain, in addition to the separate amino-acids, some lower polypeptides not attacked by trypsin.

In all, activated trypsin alone can split about half the peptide linkages in native protein. Why it leaves some simple polypeptides unattacked is not at present explained. The investigation of its action on the higher polypeptides has shown that very minute differences in the grouping of the molecule may determine whether or not the molecule is attacked by trypsin.

CONDITIONS OF TRYPTIC ACTIVITY. The contents of the duodenum are faintly acid (pH 4·5—5·1). *In vitro*, trypsin is more active in alkaline solutions. The optimum pH varies with the substrate, *e.g.* 9·7 for gelatin and 8·0 for caseinogen and peptone. However, trypsin in alkaline medium is extremely unstable, though the destructibility of trypsin, as well as its action, is largely affected by the presence of proteins or their digestion products in solution. The trypsin enters into some form of combination with the protein molecule and this protects the trypsin from the destructive action of alkali. Destruction of the enzyme nevertheless occurs in the intestine, and intestinal contents collected at the lower end of the ileum show little proteolytic activity.

THE ACTION OF PANCREATIC JUICE ON MILK. On the addition of activated pancreatic juice to milk, a clot is produced which speedily redissolves. The inactive juice has no milk-clotting action. The enzyme to which the effect is due is called *chymotrypsin*, and it closely resembles trypsin, with which it is associated. The two enzymes have, however, been separated in crystalline form.

The Action of Pancreatic Juice on Carbohydrate. The pancreatic juice, as well as a fresh extract of the pancreas itself, contains a strong amylolytic enzyme, amylase or amylopsin (optimum $p\text{H} = 7{\cdot}0$). The stages in the hydrolysis of starch brought about with pancreatic juice are exactly similar to those effected by ptyalin.

If the juice be neutralised, or faintly acidified, the process of hydrolysis goes on to the formation of glucose.‡ This is due to the presence in the juice of a second enzyme—maltase—which converts maltose into glucose. The juice in the gut is therefore able to effect the further digestion of the products of salivary digestion. On the other disaccharides pancreatic juice is without effect.

---

* KUNITZ and NORTHROP. *J. Gen. Physiol.*, 1935, **18**, 433.

† ANSON. *Ergebn. d. Enzymforschung*, 1938, **7**, 118.

‡ MELLANBY and WOOLLEY. *J. Physiol.*, 1914, **49**, 246.

The Action of Pancreatic Juice on Fats. Fresh pancreatic juice contains a powerful esterase, called *lipase*, by means of which, in the presence of water, neutral fats are broken up (*v.* p. 26) into glycerol and the corresponding fatty acids. This enzyme is active either in an alkaline, neutral or very slightly acid medium (optimum at pH = 8·0). It also hydrolyses various other esters besides fats. If the reaction be alkaline, the fatty acids produced by the lipolysis combine with the alkali present, with the formation of soaps. The enzyme may be obtained from extracts of the fresh gland, but is rapidly destroyed if active trypsin be present. It is also contained in some of the dried commercial preparations of trypsin. It is insoluble in distilled water, but easily soluble in glycerol. The velocity with which lipolysis occurs is much increased (four to five times) by the addition of bile. This adjuvant action of bile is due entirely to the bile salts. These act, firstly, as activators of lipase ; secondly, by assisting emulsification, so exposing a larger area of fat to the enzyme ; and thirdly, by helping to dissolve substances otherwise insoluble in water, as explained later.

Adaptation of Pancreatic Juice to Diet. The enzymes of the pancreatic tissue alter in quantity according to the nature of the diet, so that, *e.g.* on a high carbohydrate diet there is after a time an increased amount of amylase and a reduced amount of trypsin. After a period on a high fat diet, however, all the enzymes are reduced, though the lipase suffers least. The following table shows the type of change caused by habituation to given diets.

PANCREATIC ENZYMES OF RATS ON DIFFERENT DIETS *

Diet.	Relative Enzyme Concentration.		
	Amylase.	Trypsin.	Lipase.
Carbohydrate . .	689	9·1	333
Balanced . . .	489	20·3	358
Protein . . .	398	60·9	464
Fat	170	24·4	316

6. THE LIVER AND BILE

In the adult, the relation of the liver cells to the ducts is entirely subordinated to their relation to the blood vessels of the organ, and it requires special histological methods to make out the relations between the liver cells and the bile ducts. The circulation through the liver is described in Chapter XXXV.

Liver tissue is very rapidly removed if it becomes necrotic as, *e.g.* after occlusion of a branch of the hepatic artery or after chloroform poisoning ; it rapidly grows again after destruction or removal and regains its normal size. Even after extirpation of three-fourths of the liver the remainder soon recovers its normal size, though it is of softer texture. The liver varies greatly in size according to the state of digestion, and in fasting may be only half the normal size. The size of the liver is very delicately adjusted, and if after removal of part of the liver the animal be fasted, recovery growth ceases when the size of the fasting liver has been reached. (Rous.)

As might be expected from its structure, the secretory functions of the liver represent but a small proportion of its activities in the body. The liver is in fact the greatest chemical factory of the body, receiving by the portal vein the products of digestion as they are absorbed from the alimentary canal. It converts these into other substances, breaking them down or building them up according to the needs of the body as a whole. Hence the liver plays an important part in the metabolism of proteins, carbohydrates and fats, so that its functions will have to be dealt with in the various chapters concerned with the fate of the different foodstuffs in the body.

---

* GROSSMAN, GREENGARD and IVY. *Amer. J. Physiol.*, 1943, **138**, 676.

So-called tests of "liver function," carried out for purposes of clinical diagnosis, each examine only one aspect of the liver's many functions. Among other changes seen in hepatic derangement are a rise of plasma bilirubin, phosphatase and globulin, a fall of plasma albumin, and reduction of lævulose tolerance and of hippuric acid synthesis.

The Bile. In addition to its function in digestion, the bile must also be regarded as an excretion representing the channel by which the products of disintegration of hæmoglobin are got rid of. Hence the production of bile must be continuous, but, as a digestive fluid, it is needed in the gut only during digestion. We therefore require a continuous excretion, but a discontinuous entry of bile into the small intestine. This is secured, in the majority of animals, by the existence of the gall bladder, in which bile, secreted during the intervals between the periods of digestive activity, can be stored up. In the horse and the rat, where the gall bladder is absent, its place is taken to some extent by the great size of the bile ducts. Since the bile accumulates in the gall bladder during the whole time that digestion is not going on, and is only poured into the gut during digestion, in a fasting animal the gall bladder is distended, whereas some hours after a meal it is less full.

The common bile duct and one of the pancreatic ducts open together into the duodenum in man, the ape, horse, cat and dog ; in other animals either the pancreatic duct opens into the bile duct (sheep, rat), or else the two open separately, the pancreatic duct being the further down the duodenum (ox, pig, rabbit, guinea-pig).

The gall bladder and bile ducts have an incomplete layer of peritoneum on the outside ; the middle coat consists of plain muscle and fibrous tissue and the inner lining of tall clear columnar cells. The common bile duct runs obliquely, sometimes for a centimetre or two, through the duodenal wall before opening on its inner surface. This part forms an expanded portion, or ampulla, having a strong muscular wall, which is further reinforced near the opening by a circularly arranged layer of plain muscle fibres, the sphincter of Oddi.

The epithelium of the gall bladder and bile ducts secretes a fluid containing mucin, and also absorbs water from the bile. When the cystic duct becomes obstructed the gall bladder ultimately becomes filled with a colourless mucilaginous secretion—the so-called "white-bile" of surgeons. During its stay in the bladder the bile is rapidly concentrated by the loss of water and by the addition to it of mucin or nucleo-protein, derived from the cells lining the bladder. This is shown by comparison of the composition of bile obtained from the gall bladder with that obtained from a fistula of the bile duct. The speed with which the gall bladder concentrates the bile is remarkable, the total solids of bladder bile reaching as much as ten times that of fistula bile, and the volume showing a corresponding reduction.

The inorganic salts are mostly sodium chloride and bicarbonate, the fixed CO_2 of the bile varying from 10 to 60 volumes per cent. The reaction of liver bile in the hepatic ducts is usually slightly alkaline (pH 7·7), while in the gall bladder it is slightly acid (pH 6·8) as a rule.

ANALYSES OF BILE (HUMAN) [per cent.]

From a biliary fistula (Yeo and Herroun)		From the gall bladder (Czyhlarz *et al.*)	
Mucin and pigments	0·15	Mucin	1·64
Sodium taurocholate	0·06	Bile Pigments	0·16
Sodium glycocholate	0·16	Bile Salts	4·67
Cholesterol	0·04	Cholesterol	0·66
Lecithin		Lecithin	0·56
Inorganic salts	0·84	Inorganic Salts	0·85
Water	98·72	Water	90·57

From the point of view of digestion, the important constituents of bile are the "*bile salts*," with the lecithin and cholesterol held in solution by these salts.

The Bile Salts. Characteristic of bile are the so-called "bile salts," which give the well-known Pettenkofer reaction—a purple colour with cane sugar and strong sulphuric acid. The bile salts are sodium salts of mixtures of complex acids, called *glycocholic and taurocholic acids.* Glycocholic and taurocholic acid are both present in human bile, but in the bile of the dog and the sheep there is no glycocholic acid. These acids are compounds of the cholic acids, and can readily be resolved, by hydrolytic splitting of a peptide link by means of acids, into their proximate constituents thus :—

Glycocholic acid = glycine + cholic acid
Taurocholic acid = taurine + cholic acid.

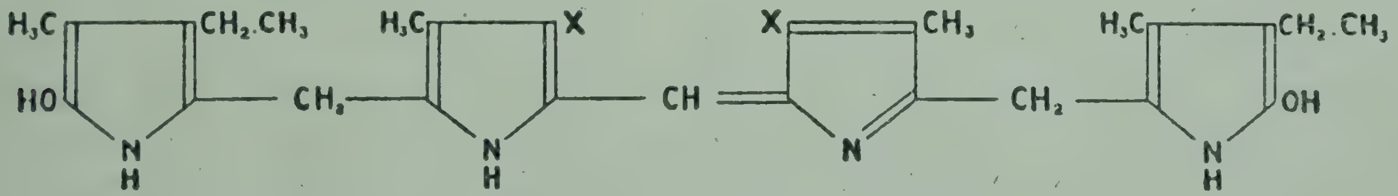
Taurine is amino-ethylsulphonic acid, $NH_2.CH_2.CH_2.SO_2.OH$, and is probably derived in the body from the metabolism of cystine.

The Cholic Acids (cholic acid, desoxycholic acid, chenodesoxycholic acid and lithocholic acid) are also often present free in the bile ; on dehydration and subsequent reduction all yield *cholanic acid*, $C_{24}H_{40}O_2$. Cholanic acid or its isomer, allo-cholanic acid, can be obtained from cholesterol by reduction, and subsequent oxidation with chromic acid. It is thus clear that the bile acids are chemically closely related to cholesterol (see formulæ, p. 28).

If cholic acid is administered it is rapidly re-excreted, and largely in the conjugated form if the liver is normal, and if glycine or taurine is also given or made available, as when a liberal protein diet is given. Although it is probable that the conjugation of the bile salts is effected in the liver, there is no conclusive evidence for this, and very little for the supposition that, in the body itself, cholic acid is actually derived from cholesterol.

Many substances, *e.g.* fatty acids, higher alcohols, etc., form loose compounds, called *choleic acids*, with desoxycholic acid, and as these compounds are soluble in water, the bile salts are capable of bringing into aqueous solution many normally insoluble substances. This is called "hydrotropic action."

The Bile Pigments, *bilirubin* and *biliverdin*, are by-products, formed by the cells of the reticulo-endothelial system in various parts of the body from hæmoglobin of the blood (p. 522). Injection of hæmoglobin into the blood stream, or extensive destruction of blood in the body, lead to an increased formation of bile pigment. The pigment thus formed circulates in small amount in the blood plasma from which it is excreted by the liver into the bile. The bile pigments pass into the intestine, and are there converted by bacterial reduction into stercobilin which is excreted for the most part with the fæces, a small proportion being absorbed into the blood and turned out as stercobilin in the bile or as urobilin in the urine. When the bile of a dog is all drained off by a biliary fistula, stercobilin disappears from the fæces and bile ; if bile is given by mouth it reappears. When absorption of stercobilin is excessive, urobilin appears in the urine.



Urobilin. ($X = -CH_2 \cdot CH_2 \cdot COOH.$)

The Secretion of Bile. It is important to distinguish between the formation of the bile in the liver, which is a continuous process, and the expulsion of the bile, either direct or from the gall-bladder, into the intestine, which is discontinuous and principally dependent on the taking of food. The time-relations and other phenomena of the secretion of bile can be learnt from animals or men in which, by operation, the bile is conducted to the outside by means of a permanent fistula. Very little reliable information can, however, be obtained direct from men, and most of our information has been obtained from dogs. Of the various types of operation suited for the study of the bile flow, the best is probably that called "altercursive intubation," is illustrated in Fig. 570.

The secretion of the bile is independent of the action of the nervous system on the liver, since complete denervation of the liver leaves it unaltered and does not affect the influence of food on bile secretion.[*] Bile formation also occurs in an excised perfused liver. Though the bile is always being secreted, the amount and concentration are both lowered during fasting. It is also reduced if the bile as it is formed is collected instead of being allowed to enter the duodenum.

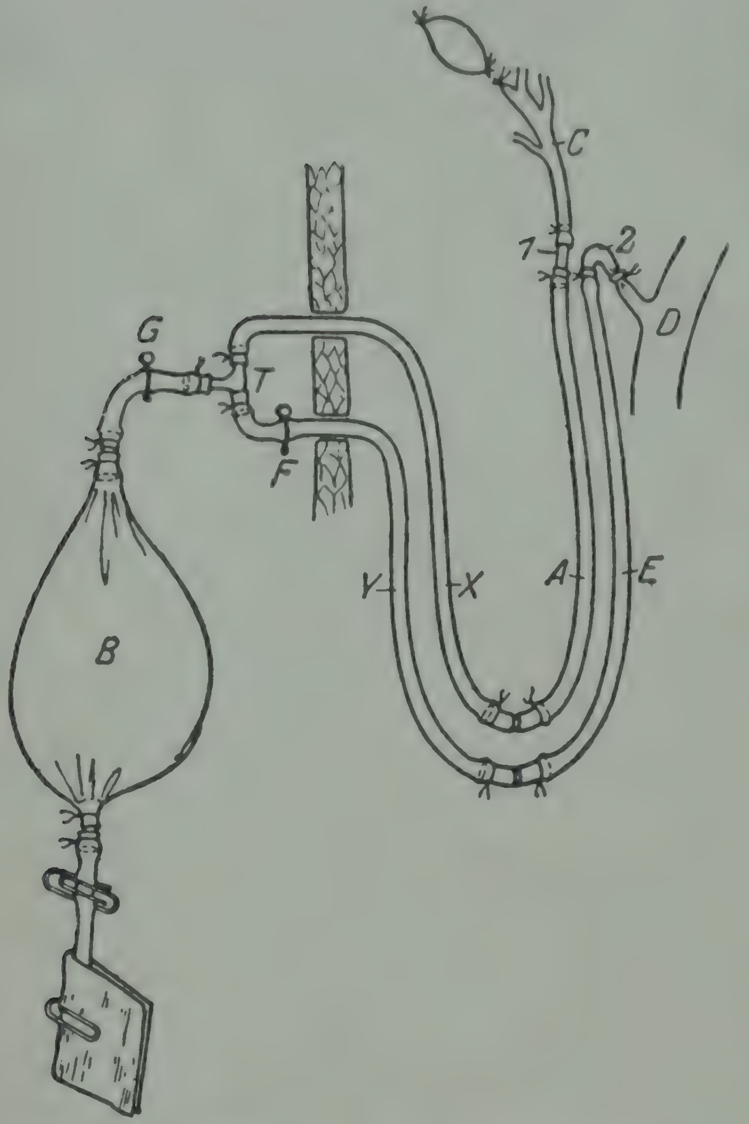


FIG. 570. Schema of the operation for "altercursive intubation." The bile can flow normally from *C* along the tubes *A*, *X*, *T*, *Y*, *E*, into the natural opening in the duodenum. When desired, *F* is closed and *G* opened, so that bile collects in the rubber bag *B*. A manometer can be connected to *X* to register secretion pressure, or a bile-filled manometer connected to *Y* to measure the resistance to flow of bile into the duodenum. The gall-bladder is excluded by ligatures. (McMaster and Elman, *J. Exp. Med.*, 1925, **41**, 513.)

If the hepatic duct be obstructed, as by attaching to it a manometer (Fig. 570) filled with bile, it is found that the flow continues at a constant rate, so that the pressure rises uniformly, until a pressure of about 300 mm. of bile is reached, at which point it abruptly ceases. At this level reabsorption of the bile into the lymph and/or blood takes place.

The rate of bile secretion in man is about 15 c.c. per kg. body weight per day, or, say, about 800–1,000 c.c. for average adults ; in dogs it varies in different individuals between about 5 c.c. and 20 c.c. per kg. per day. When food, especially meat or fat, is taken, the rate of secretion begins to increase in about an hour and reaches a maximum in two to five hours after the taking of food, then declining to the fasting rate. Carbohydrate food causes no increase. The secretion of the bile, after meat or fat, thus runs roughly parallel to that of the pancreatic juice.

Various other substances cause an increased bile secretion when introduced into the stomach or duodenum, even after the division of all the nervous connections of the liver. The most potent of these is bile, or the bile acids,[†] which also increase the total solids of the bile formed. The active substance is cholic acid, which acts similarly when given by intravenous injection. The introduction of dilute acids into the duodenum, or the intravenous injection of secretin [‡] also greatly augments the bile flow. Probably an alteration in the blood flow to the liver is an important factor in modifying the rate of bile secretion.

Bile Storage and Expulsion. The alterations occurring in the size of the gall-bladder can be most readily observed clinically by the administration of tetra-iodo-phenol-phthalein, which is excreted in the bile and is opaque to the X-rays, so that the outline of the gall-bladder can be seen on X-ray examination (Graham-Cole test).

---

[*] Lundberg. *Amer. J. Physiol.*, 1931, **98**, 602 ; Hillyard. *Ibid.*, 612.

[†] J. Mellanby. *J. Physiol.*, 1928, **64**, 331.

[‡] Still, McBean and Ries. *Amer. J. Phys.*, 1932, **99**, 94.

The storage of bile in the gall-bladder and its expulsion therefrom are best investigated by simultaneously studying the pressure of bile in the common duct and the resistance to its entry into the duodenum. This may be done by the triple intubation operation of McMaster and Elman,* shown in Fig. 571. The sphincter of Oddi is competent to close the entrance of the bile duct into the duodenum when the pressure of bile is low (below 150 mm. H_2O), and its action is further aided by contraction of the muscular wall of the duodenum itself.† After a fast of twenty-four hours, or on a carbohydrate diet, the resistance to the entry of bile into the duodenum is high, sometimes as great as 300 mm. H_2O. Hence the bile as it is secreted by the liver is diverted into the gall-bladder, and, since it is there very rapidly concentrated by loss of water, the storage of bile can continue for days without there being a great rise of pressure in the bile ducts, in which it is usually about 100 mm. H_2O during fasting. There is thus no entry of bile into the duodenum, since a pressure over 300 mm. would be necessary to force the sphincter.

When food is seen or eaten, there is at once a lowering of the sphincter resistance and a contraction of the gall-bladder great enough momentarily to raise the pressure in the duct system to about 200 mm. H_2O. This is probably assisted by a contraction of the ampullary part of the common duct. A gush of bile into the duodenum follows. Afterwards the resistance suddenly rises again, but in about half an hour another rise of pressure with relaxation of the sphincter occurs, and thereafter at fairly frequent intervals for the next few hours bile is expelled into the duodenum.

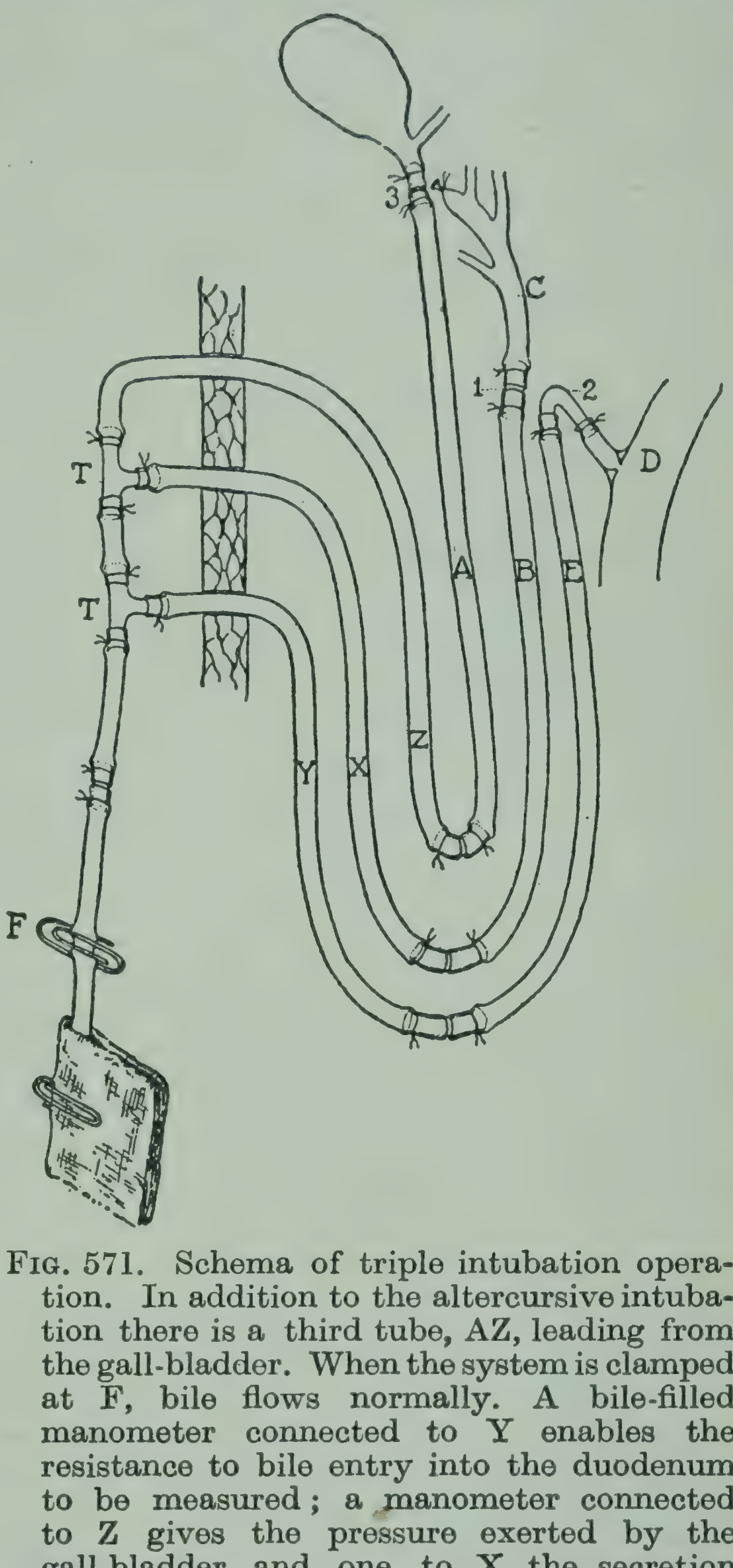


Fig. 571. Schema of triple intubation operation. In addition to the altercursive intubation there is a third tube, AZ, leading from the gall-bladder. When the system is clamped at F, bile flows normally. A bile-filled manometer connected to Y enables the resistance to bile entry into the duodenum to be measured; a manometer connected to Z gives the pressure exerted by the gall-bladder, and one to X the secretion pressure of the bile. (McMaster and Elman.)

The control of the gall-bladder seems to be a twofold one—by a reflex and by a hormone. The muscular wall of the gall-bladder shows rhythmic contractions, and is under the control of the vagus and the sympathetic,‡ the former conveying motor and the latter inhibitory impulses, while the sphincter of Oddi receives branches from the same sources, but having the opposite functions. It seems certain that the sight of food, or the act of eating it, causes a reflex contraction of the gall-bladder, with reciprocal relaxation of the sphincter of Oddi. It has also been supposed that the entry of acid into the duodenum provokes a similar reflex,* and there is clear evidence that under these circumstances the resistance of the sphincter is reduced.

---

* McMaster and Elman. *J. Exp. Med.*, 1926, **44**, 184; Elman and McMaster, *ibid.*, 1926, **44**, 151.

† Lueth. *Amer. J. Physiol.*, 1931, **99**, 237.

‡ Bainbridge and Dale. *J. Phys.*, 1905, **33**, 138; Westphal, *Z. f. klin. Med.*, 1923, **96**, 52.

That much of the activity of the gall-bladder is due to the action of a hormone liberated into the blood by the presence of certain substances in the duodenum seems equally certain. Acids, fat and egg-yolk are the most active substances, and cause contraction even of the denervated gall-bladder, when they are introduced into either the intact intestine, or a transplanted loop. Ivy and Oldberg † have extracted from the duodenal mucosa a secretin-like substance, "*cholecystokinin*," which stimulates contraction of the gall-bladder when injected.

Introduction of magnesium sulphate into the duodenum causes relaxation of the sphincter and expulsion of bile. Morphine causes strong spasm of the sphincter.

It is of interest to note that the expulsion of bile into the duodenum serves a twofold function. First, it promotes pancreatic secretion by aiding in the transference of secretin into the blood, and secondly, it augments the formation of the bile itself.

Entero-hepatic Circulation. Although bile is regularly passed into the intestine, very little bile salts are to be found in the fæces. This is because they are reabsorbed from the intestine into the portal blood, whence they are picked up and excreted again by the liver : this is called the entero-hepatic circulation of the bile salts. Hence, if a biliary fistula is made, so that the bile is lost from the body, the cholic acid derivatives of the bile rapidly decline, but they finally reach a steady level at about 100 mg./kg/day in dogs, or about 1·5 g./day for man. This means that new formation of cholic acid occurs at that rate—the so-called " endogenous production." If now, in the fistula dog, we return bile of known cholic acid content to the duodenum, the output of cholic acid by the bile rises by an amount equal to that administered, minus about 12 per cent., whatever the amount of cholic acid, within limits, given.‡ If we thereafter regularly return all the bile produced, the daily output of bile acids steadily subsides (or rises), until it sooner or later reaches a new steady level, at which synthesis exactly balances loss. This level depends on the diet, but can be about 1 g./kg./day for the dog. One of the factors which brings about this homoiostatic control is probably the effect of bile salts in increasing the rate of secretion of bile, so quickening up the entero-hepatic circulation, and increasing the loss (= 12 per cent. of turnover), and *vice versa*. Owing to the fact of their ready excretion by the liver cells, bile salts are the most reliable cholagogues with which we are acquainted. By this circulation of salts between liver and intestine, the synthetic work of the liver in the production of the bile salts is reduced to a minimum, and it has only to replace such of the bile salts as undergo destruction in the alimentary canal under the influence of micro-organisms, or as are lost to the organism by passing out in the fæces as a gummy amorphous substance known as *dyslysin*.

There appears also to be a similar circulation of the bile pigments. Thus, if in a dog with altercursive intubation the bile is collected in the bag for a time, the pigment content of the later samples secreted becomes reduced. If the bile be now returned to the duodenum, the pigment content of the subsequently secreted samples of bile is increased again. If the dog be fed with the bile of an herbivorous animal, the characteristic herbivorous pigment appears in the dog's bile. The presence of bile in the intestine seems to be essential to health. If the bile is all led away, bone defects are liable to occur, and a serious state of anæmia sets in ; the animal is restored to health again by being fed with bile, or with blood or blood pigment.

The Digestive Functions of the Bile. The digestive functions of bile are due, not to the presence of enzymes, but to the bile salts. Most important is the part played in the digestion and absorption of fats. The fat-splitting action of pancreatic juice is at least trebled by the addition of bile, whether boiled or unboiled.

---

* OKADA. *J. Physiol.*, 1915, **50,** 42.

† IVY and OLDBERG. *Amer. J. Physiol.*, 1928, **86,** 599 ; LUETH, IVY and KLOSTER, *ibid.*, 1929, **91,** 329.

‡ BERMAN *et al.* *Amer. J. Physiol.*, 1941, **131,** 776.

In the first place, the bile salts act as powerful emulsifying agents, as explained below.

Further, the bile salts are activators of pancreatic lipase, even when this acts on substances in solution. It is possible that this activating effect is due to the bile salts in some way bringing the enzyme and substrate into sufficiently close contact with one another. Lipase is insoluble in water, but soluble in fatty media. An important digestive function of bile, therefore, lies in the power of bile salts of serving as a vehicle for the suspension and solution of the interacting fats, fatty acids, and fat-splitting enzyme. This vehicular function is probably due to hydrotropic action, and plays an important part in the absorption of fats and cholesterol. These pass, with bile salts, through the striated basilar membrane bounding the intestinal side of the epithelium, and the bile salts are then released and carried by the portal vein to the liver.

The action of pepsin is checked by bile owing to its inclusion in a precipitate which is thrown down.

Emulsification of Fat. In the course of its digestion, triglyceride fat is very finely emulsified in the intestine, the particles being negatively charged, and the emulsion being stable over a considerable *p*H range, say 6·0–8·5. So many possible emulsifying agents (*v.* p. 27) are present in the intestine that attempts to explain emulsification suffer from a confusion of choice. Thus soap, fatty acid, monoglyceride, cholesterol, lecithin, bile salts, mucin and other proteins, can all, under certain conditions, act as emulsifying agents.

Trial, *in vitro*, of possible substances alone or in varied double or triple combinations indicate that the requisite criteria are satisfied by a triple combination of monoglyceride, fatty acid and bile salts. In digestion the monoglyceride and fatty acid are formed by lipolysis of some triglyceride, and then aid in emulsifying the remainder of it. The system of three agents ensures stability over wide *p*H ranges ; fatty acid supplies the charge on the monoglyceride-stabilised particle at the more alkaline *p*H ranges, while the fatty acid/bile salt complex provides the charge when the ionisation of the fatty acid is depressed in more acid media.*

Jaundice. If the bile ducts are ligatured, or obstructed in any way, so as to prevent the entry of bile into the intestine, the bile is reabsorbed and passes by the lymphatics into the blood stream. The bile pigment circulating in the blood permeates the various tissues of the body, giving them a yellow colour ; bile pigments and bile salts are also excreted with the urine. This condition is known as jaundice. Certain types of jaundice have long been known in which no apparent obstruction of the bile ducts was present. It was therefore assumed that in these cases the bile pigment was set free in the blood, apart from the activity of the liver. Pathologists thus distinguished two kinds of jaundice, hepatogenous and hæmatogenous. The existence of the latter form was long disputed, although it was known that bilirubin could be formed as a product of the disintegration of hæmoglobin in old blood clots. F. C. Mann succeeded in totally extirpating the liver, and, by means of injections of glucose, in keeping the animals (dogs) alive for thirty-six hours. He found that, six hours after the extirpation, bile pigment begins to accumulate in the blood and in the tissues, showing that this substance can be formed as the result of the disintegration of hæmoglobin without the co-operation of the liver, thus justifying the long-disputed division of jaundice into the two varieties of hepatogenous and hæmatogenous.

This conversion of hæmoglobin into bilirubin occurs in the scattered intravascular macrophages in the liver, spleen and bone marrow, all of which belong to the reticulo-endothelial system.

For Reference

Horrall, O. H. (1931). The Toxicity of Bile. *Physiol. Rev.*, **11,** 122.
Ivy (1934). The Physiology of the Gall Bladder. *Physiol. Rev.*, **14,** 1.
Josephson (1941). The Circulation of the Bile Acids in Connection with their Production, Conjugation and Excretion. *Physiol. Rev.*, **21,** 463.
Mann, F. C. (1924). The Functions of the Gall Bladder. *Physiol. Rev.*, **4,** 251.

---

* Elkes *et al.* *J. Physiol.*, 1944, **103,** 6.P.

Rich, A. R. (1925). The Formation of Bile Pigment. *Physiol. Rev.*, **5**, 182.
Schmidt, C. L. A. (1927). The Extra-hepatic Functions of Bile. *Physiol Rev.*, **7**, 129.
Snyder (1942). Recent Advances in Knowledge of the Liver. *Physiol. Rev*, **22**, 54.
Sobotka (1937). Physiological Chemistry of the Bile. (Baltimore.)
Whipple, G. H. (1922). The Origin and Significance of the Constituents of the Bile. *Physiol. Rev.*, **2**, 440.

7. THE JUICES OF THE SMALL INTESTINE

The juice of the small intestine consists of the secretion of Brunner's glands, together with the *succus entericus*, or intestinal juice proper.

The glands of Brunner extend only over about an inch at the very beginning of the duodenum in the carnivora, but in the herbivora they may be found occupying the upper six inches of the intestine.

The duodenal secretion is rich in mucus, and is alkaline. There is no secretion during fasting, but juice begins to be formed within a few minutes of the taking of food. The natural flow occurs at normal rates after section of all nerves, and also in the transplanted duodenum.* The glands appear to contain a peptidase and an esterase,† but the only significant enzymes in their secretion are amylase and enterokinase. One function of the juice may be, by its mucin, to protect the mucosa from the gastric juice.

The intestinal juice proper is probably secreted by the Lieberkühn's follicles but the villi may also contribute to it. There are differences between the upper, middle and lower portions of the gut ; secretion is better marked in the upper portions, absorption predominates in the lower sections.

Secretion in the small intestine may be studied by isolating loops, in an anæsthetised animal, or better, by establishing permanent fistulæ, the best type being the Thiry-Vella. This consists in separating a loop of intestine, and restoring the continuity of the gut from which the loop had been excised. Both ends of the loop are left open and brought into the abdominal wound. In such a fistula it is easy to introduce substances into the upper end and determine the flow of juice from the lower end, the constant emptying of the loop being provided for by the peristaltic contractions of its muscular coat. Intestinal fistulæ resulting from war injuries have been studied in man by Savitsch, and have given results confirming those made on dogs. Observations can also be made by intestinal intubation in normal man.‡

In animals with intestinal fistulæ, no qualitative difference has been recorded between the upper and lower ends of the gut. Fasting animals show no secretion, or merely a scanty periodic flow about every two hours, from an empty loop. Administration of food is also without effect on the isolated loop, so long as this remains empty. Vomiting excites a reflex flow, through efferent vagal impulses.§ An important factor in exciting a flow of juice is local mechanical stimulation : thus the introduction of a drainage tube into the isolated loop is followed by a steady flow of succus entericus, whether the animal be fasting or fed. The flow of juice thus initiated is greatly modified by feeding the animal ; thus, a meal of fat will often double the rate of secretion.

It seems probable that the normal flow of intestinal juice is provoked locally by the presence, in the intestine, of its normal contents. One factor is the mechanical stimulus of the contents ; the other is probably a local chemical stimulation by various products of digestion.

The response to local mechanical or chemical excitation is apparently brought about through the local nerve net (Meissner's plexus), since section of the vagus does not abolish the reaction. But if nervous influence normally plays any part in causing secretion, it is probably effected through the vagi, since the rate of the secretion can be augmented by stimulation of the vagus, though only after a very long latent period (one hour), and the juice obtained, unlike that resulting from local mechanical stimuli, is very rich in enzymes. (Savitsch.)

As might be expected, cholinergic substances introduced into the blood stream, *e.g.* acetyl choline, eserine, or pilocarpine, produce secretion of the juice.

Certain facts seem to speak for an inhibitory action of the central nervous system on intestinal secretion. Thus it has been observed that extirpation of the nerve plexuses of the abdomen, or section of the splanchnic nerve, causes diarrhœa for two or three days. This might be determined by removal of inhibitory impulses normally arriving at either the muscle wall or the intestinal glands.

---

* Florey and Harding. *Proc. R. S.*, 1935, **117B**, 68 ; *Q. J. Exp. Phys.*, 1935, **25**, 329.

† Linderstrom-Lang and Holter. *C. R. Lab. Carlsberg*, 1925, **20**, 42.

‡ Owles. *Clin. Sci.*, 1937, **3**, 1.

§ Gregory. *J. Physiol.*, 1947, **106**, 95.

There have been claims to have demonstrated the presence in the intestinal mucosa of a specific secretory hormone, called *enterocrinin*. Injection of crude (but not of pure) secretin causes a secretion of intestinal juice which may, of course, be attributed to the presence in it of "enterocrinin." It is also possible that general humoral mechanisms may play a part, though probably only a subsidiary one, in bringing about normal secretion of succus entericus.

Characters of Intestinal Juice. The intestinal juice obtained from a permanent fistula has a specific gravity of about 1,010. It is generally turbid from the presence in it of migrated leucocytes, disintegrated epithelial cells and clumps of mucus. It contains about 1·5 per cent. total solids, of which 0·8 per cent. are inorganic, chiefly sodium bicarbonate and sodium chloride. Its *p*H varies from 6·3 to 8·6 according to site, and, especially, to the amount of CO_2 lost in collection. It is less buffered than pancreatic juice. The organic matter includes a number of enzymes, but it is an open question how many of those nominated are actually present in the juice secreted under normal conditions. Some of the claims rest upon the examination of bacterially contaminated scrapings of the mucosa, others on the turbid and cell-rich juice mixed with desquamations, obtained by collection from fistulæ. While it is probable that the normal juice is always mixed with some shed epithelial cells, the relative number of these may be small compared with those set free in a fistulated loop exposed to a varying degree of mechanical insult. The fair thing seems to be to consider the enzymes found, apart from bacteria, in (*a*) centrifuged fresh juice, (*b*) fistula juice as collected, and (*c*) mucosa scrapings.

(*a*) Centrifuged fresh juice always contains *enterokinase*, the activator of pancreatic trypsinogen (p. 879), and *amylase*.*

(*b*) The turbid juice, as collected, usually contains various *peptidases*, *lactase*, *maltase*, *invertase* and *lipase*. The peptidases, formerly considered as a single enzyme called erepsin,† are distinguished by the fact that they have no power of acting on native proteins, peptones or proteoses, but only upon the polypeptides, which are rapidly split (optimal *p*H = 7·8). The juice contains several different peptidases which act on peptides of different structure, a dipeptidase acting on dipeptides, etc., and convert them into amino-acids.

(*c*) Scrapings of the mucosa contain all the enzymes mentioned above, and also, often or usually, *proteases*, *nuclease*, *phosphatase*, and *arginase*.

We may suppose that the enzymes discussed above exercise their functions in digestion in three ways. The soluble enzymes present in the juice, viz. enterokinase and amylase, exert their proper actions on trypsinogen and on starch respectively. The shed cells also present to a varying extent in the juice break down and set free their insoluble enzymes into the lumen of the gut, and these also exert their effects in breaking down polypeptides, disaccharides, or fats. The insoluble enzymes which are retained in the normal cells of the intestinal mucosa exert their actions on substances in transit through the epithelium, in the normal course of absorption, *i.e.* they are proper endo-enzymes.

The functions of the intestinal enzymes may therefore be said to be, first, the activation of the pancreatic trypsinogen, second, the further breakdown of all the constituents of the chyme, and third, the action, as, *e.g.* by phosphatase, upon substances in the course of absorption. In addition, the intestinal juice ensures a continual supply of water to the intestinal contents, thus facilitating solution and absorption of the digestion products.

For Reference

FLOREY, WRIGHT and JENNINGS (1941). The Secretions of the Intestine. *Physiol. Rev.*, **21**, 36.

---

* WRIGHT *et al.* *Q. J. Exp. Physiol.*, 1940, **30,** 73.

† COHNHEIM. *Z. f. physiol. Chem.*, 1901, **33,** 451.

8. MOVEMENTS OF THE SMALL INTESTINE

The movements of the intestines can be investigated either by direct observation of the gut or by means of the barium meal and X-ray examination.* In using the former method it is essential to exclude the many reflex influences which can play on the intestine, either by extirpation of the whole of the nerve plexuses of the abdomen, by division of the splanchnic nerves, or by destruction of the spinal cord below the mid-dorsal region. It is questionable, however, whether the movements seen under these conditions are the same as those which occur in the intact animal.

Three types of movement have been described in the small intestine, namely: (1) Pendular movements; (2) segmenting movements; (3) peristalsis. The movements may be seen in an animal treated as described above, when its abdomen is opened in a bath of warm normal salt solution.

The *pendular movements* probably only occur in the opened abdomen, where they appear as writhing movements of individual loops of the intestine. Bayliss and Starling appear to have described under this name the regular movements now called segmenting, and it is doubtful whether the pendular movements are anything more than the expression of segmenting contractions freed from the confinement of their normal surroundings.

The ***segmenting movements*** are the most constant and fundamental. They may be recorded by inserting into the intestine a balloon filled with air and connected with a piston recorder. The constriction which accompanies each contraction is then seen to be most pronounced at the middle of the balloon, i.e. in the zone of greatest distension, and the amplitude of the contractions is augmented by increasing the tension on the walls of the gut. These movements are unaffected by the direct application of drugs, such as nicotine or cocaine, which we might expect to paralyse any local nervous structures in the intestinal wall; they occur equally well in the isolated intestine, from which facts Bayliss and Starling † concluded that these rhythmic contractions were myogenic, and were propagated from muscle fibre to muscle fibre. This is confirmed by the observation of Gunn and Underhill ‡ that the contractions continue in strips of muscle completely freed from any remains of the nerve plexus. These contractions occur with considerable regularity at a frequency which varies inversely with their distance from the stomach, and in the ileum is of the order of twelve a minute. The different frequencies at different levels is due, according to Alvarez, to the existence of a "metabolic gradient" from

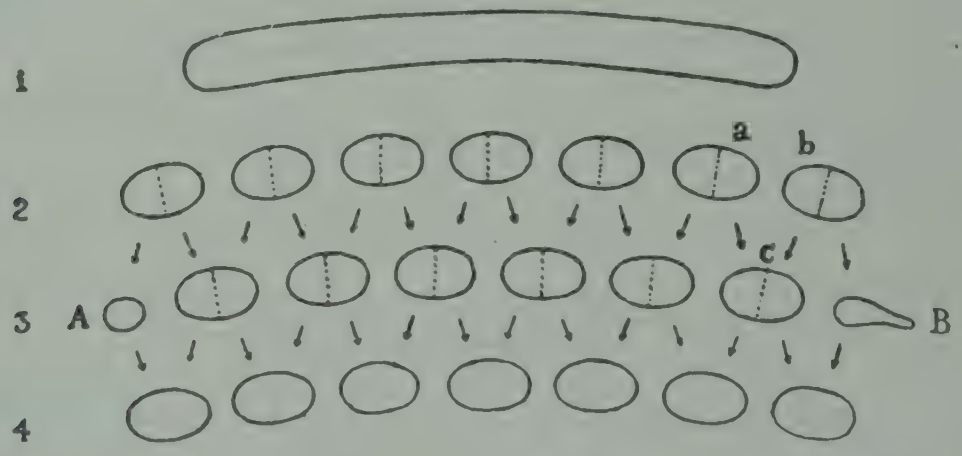


FIG. 572. Diagram of the "Segmentation" Movements of the Intestines as observed by the X-Rays, after Administration of Bismuth. (CANNON.)

1. A continuous column, intestinal movements being absent.
2. The column broken up into segments.
3. Five seconds later, each segment divided into two, the halves joining the corresponding halves of adjacent segments.
4. Condition (2) repeated five seconds later.

---

* CANNON. *Amer. J. Physiol.*, 1902, **6**, 250.

† BAYLISS and STARLING. *J. Physiol.*, 1899, **24**, 99; 1901, **26**, 125.

‡ GUNN and UNDERHILL. *Quart. J. Exp. Phys.*, 1914, **8**, 275.

above downwards, the upper portions having a higher rate of metabolism and being also more excitable than the lower parts.

These segmentations must cause a thorough mixing of the contents of the gut with the digestive fluids. On examining under the X-rays the intestine of a cat, which some hours previously has taken a large meal of bread and milk mixed with bismuth, a length of gut may be seen in which the food contents form a continuous column. Suddenly movements occur in this column, which is split into a number of equal segments. Within a few seconds each of these segments is halved, the corresponding halves of adjacent segments uniting. Again contractions recur in the original positions, dividing the newly formed segments of contents and re-forming the segments in the same position as at first (Fig. 572). In this way every particle of food is brought successively into intimate contact with the intestinal wall. These movements have not a translatory effect, and a column of food may remain at the same level in the gut for a considerable time.

The ***peristaltic waves*** cause onward progress, possibly with a spiral rotation,* of the contents of the intestine. The classical description of peristalsis is that given by Bayliss and Starling, who described it as a wave moving slowly—generally at about 2 cm. a minute—involving contraction of the gut above the food mass, and relaxation below it. If a balloon be inserted in the lumen of the exposed gut, it will be found that pinching the gut *above* the balloon causes an immediate relaxation of the muscular wall in the neighbourhood of the balloon. On the other hand, pinching the gut half an inch below the situation of the balloon causes a strong continued contraction to occur at the level of the balloon itself (Fig. 573). Stimulation at any portion of the gut causes contraction above the point stimulated and relaxation below it (the " myenteric reflex "). The same effect is produced by introduction of a bolus of food, especially if it be large or have a direct irritating effect on the wall of the gut (Fig. 574). In this case the contraction above and the inhibition below cause an onward movement of the bolus, which travels slowly down the whole length of the gut until it passes through the ileocæcal opening into the large intestine. According to Alvarez, the direction of the peristaltic wave is determined by the direction of the metabolic gradient of the intestine. The peristaltic contraction involves the co-operation of the local nervous system, viz. Auerbach's plexus. Whereas in the œsophagus it is the central nervous system which co-ordinates, the peristaltic contractions in the small intestine continue after severance of all connection with the brain and spinal cord. On the other hand, they are absolutely abolished by painting the intestine with nicotine or with cocaine.

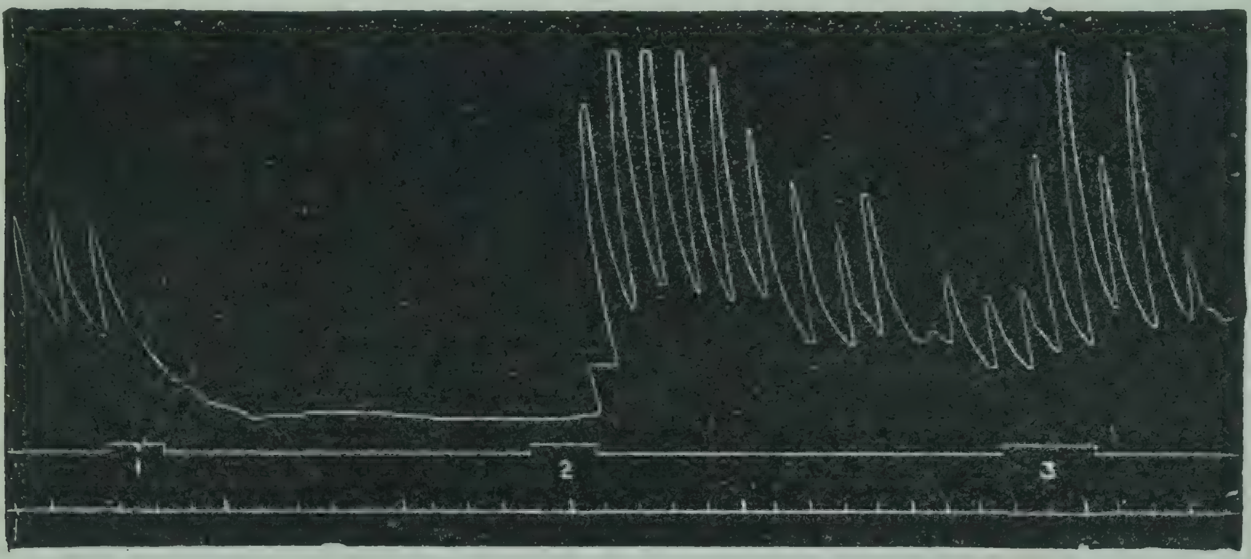


FIG. 573. Intestinal Contractions (Balloon Method).

In this dog all the abdominal ganglia had been excised, and both vagi cut. Showing propagated effects of mechanical stimulation above and below the balloon. (1) pinch above, (2) pinch below, (3) pinch below balloon.

---

* FRANKLIN and MAHER-LOUGHNAN. *J. Physiol.*, 1938, **94**, 426.

According to Alvarez, the peristaltic wave, as described by Bayliss and Starling, is abnormal, and in his opinion true peristalsis consists of a series of "rushes" traversing the whole length of the small intestine, travelling much more rapidly (2 to 25 cm. per second), and in which the constriction is less intense and not preceded by a wave of relaxation. Probably both forms occur.

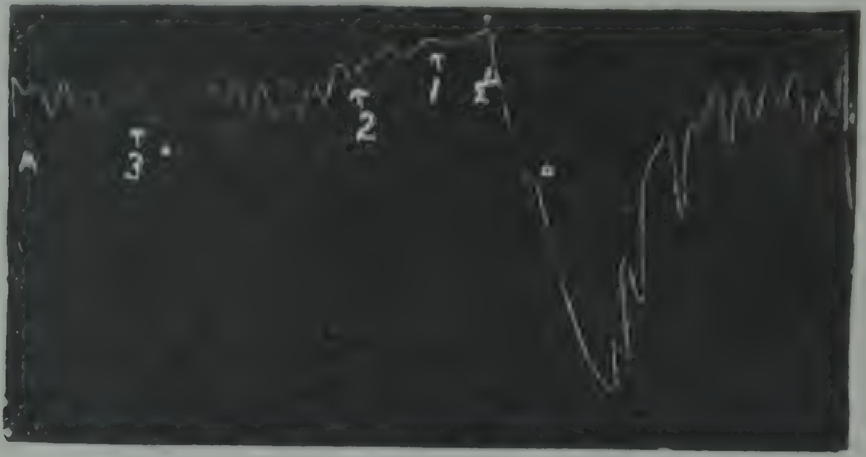


Fig. 574. Passage of Bolus. Contractions of Longitudinal Coat (Enterograph).

The bolus (of soap and cotton wool) was inserted into the intestine four inches above the recorded spot at A. The figures below the tracing indicate the distance of the middle of the bolus from the recording levers. As the bolus arrives two inches above the levers, there is cessation of the rhythmic contractions and inhibition of the tone of the muscle. This is followed, as the bolus is forced past, by a strong contraction in the rear of the bolus.

Anti-peristalsis is never observed in the lower small intestine. Mall showed that, if a short length of gut be cut out and reinserted in the opposite direction, a species of partial obstruction results, in consequence of the fact that the peristaltic waves, started above the zone of operation, cannot travel downwards over the reversed length of gut. The intestine above this level therefore becomes dilated. If, however, the reactions of the local nervous system be paralysed or inhibited, a reflux of intestinal contents is quite possible, since the contractions excited at any spot by local stimulation of the muscle have the effect of driving the food either upwards or downwards; the direction of movement of the food will be that of least resistance.

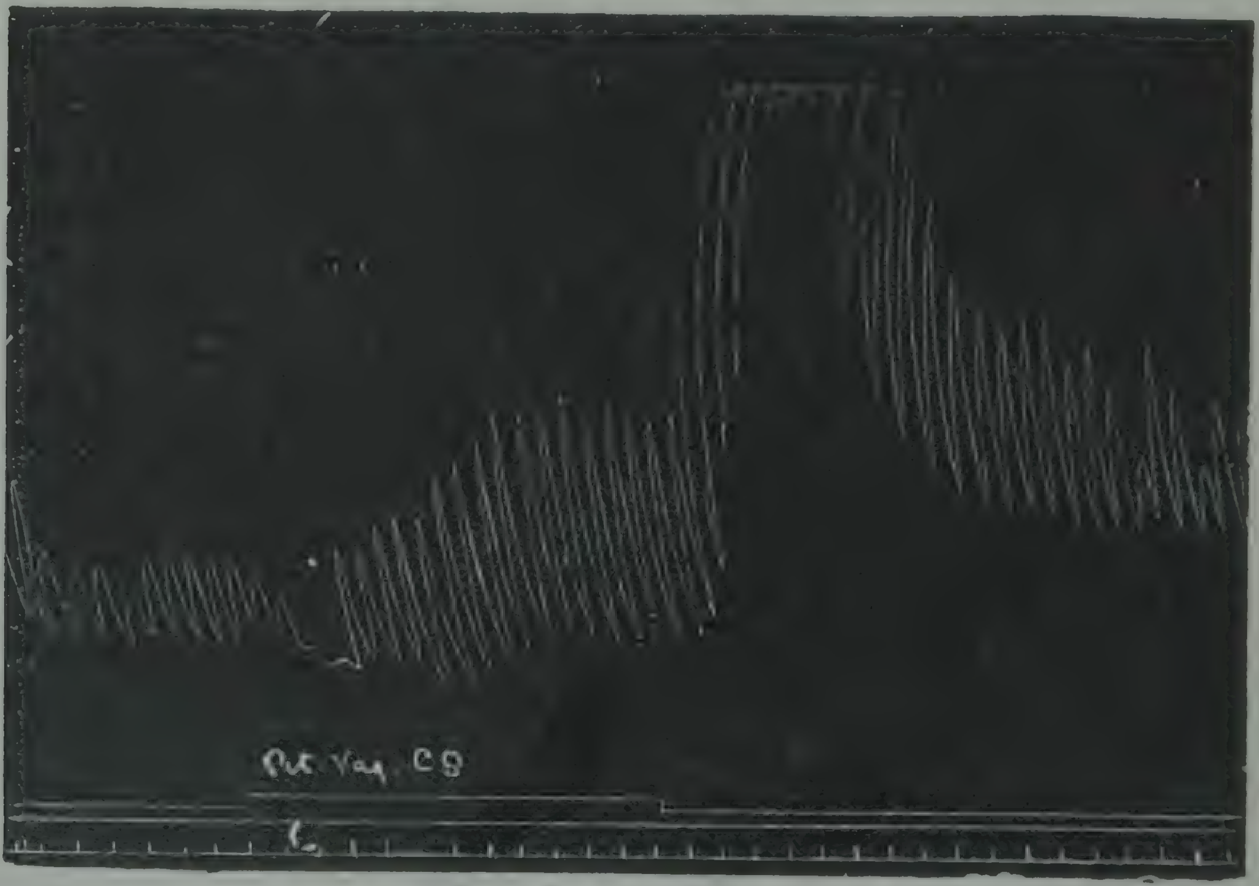


Fig. 575. Effect of Stimulation of Right Vagus on Intestinal Contractions.

The movements of the small intestine are also affected by the central nervous system. Stimulation of the vagus causes an initial inhibition of the whole small intestine, followed by increased irritability and increased contractions (Fig. 575). The ingestion of pleasant foods also increases intestinal tonus, probably by augmenting vagal impulses, while section of the vagus lowers tone. On the other hand, stimulation of the splanchnic nerves, or injection of adrenaline, causes complete relaxation of both coats of the small gut (Fig. 576). It seems that the splanchnics normally exercise a tonic inhibitory influence on the intestinal movements, which can be increased reflexly, and notably from the intestine itself: both afferent and efferent paths are in the splanchnics.* The relaxed condition of the gut, called paralytic ileus, which obtains in many abdominal affections, is probably due to reflex inhibition through the splanchnic nerves.

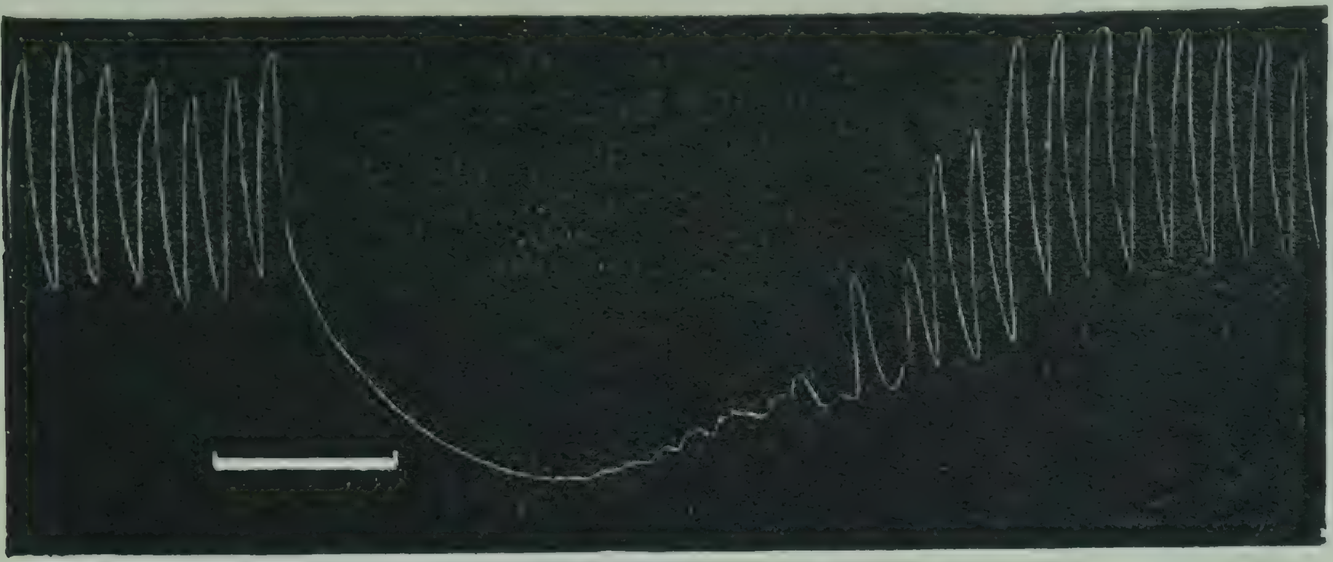


Fig. 576. Excitation of both Splanchnic Nerves. Balloon Method. Intestine returned to Abdomen.

The movements of the duodenum show some special features, viz. contraction of the first part (duodenal cap), peristalsis of the second and third parts, antiperistalsis of the second and third parts, and segmenting movements. Normally the contents are passed to and fro a few times in the duodenum before passing on to the jejunum, but once they have passed on they do not return. Antiperistalsis is most definite in the second part of the duodenum, and leads to periodic filling of the duodenal cap, or even to regurgitation into the stomach. Very rapid mixing of the contents with the digestive juices results from this to-and-fro movement.†

In man, the chyme remaining from the digestion of a meal begins to arrive at the lower end of the ileum in about three and a half hours after the meal. Here it accumulates for a while, during which time active segmentation movements occur, and further absorption of water takes place. Peristalsis is generally feeble in this part of the bowel, but is reflexly increased during or just after a meal. It is at this time, *i.e.* when the next meal is taken, that the contents are passed on by occasional peristaltic contractions through the lower end of the ileum and into the cæcum. The lowest 2 cm. of the ileum present a distinct thickening of the circular muscular coat, forming the *ileocolic sphincter*. This sphincter relaxes in front of a peristaltic wave and so allows the passage of food into the colon. On the other hand, it contracts as a rule against any regurgitation which might be caused by contractions in the colon.‡ Although thus falling into line with the rest of the muscular coat as concerns its reaction to stimuli arising in the gut above or below, it presents a marked contrast to the rest of the gut in its relation to the central nervous system. It is unaffected by stimulation of the vagus. Stimulation of the splanchnic, however, which causes complete relaxation of the lower part of the ileum with the rest of the small intestine, produces a strong contraction of the plain muscle forming the ileocolic sphincter.* According to Hurst,† an important function of the ileocolic sphincter is, by its tonic contraction, to prevent the contents of the ileum passing too rapidly into the cæcum, so as to allow more time for digestion and absorption of foodstuffs to be carried to completion in the ileum.

---

* Chang and Hsu. *Q. J. Exp. Physiol.*, 1942, **31**, 311.

† Bolton and Salmond. *Lancet*, 1927, 1230.

‡ Rutherford. "The Ileo-cæcal Valve," 1914. (London: Lewis.)

For Reference

Alvarez (1940). An Introduction to Gastro-Enterology. (Heinemann.)
Barclay (1936). The Digestive Tract (Radiological Study). (Cambridge.)

9. FUNCTIONS OF THE LARGE INTESTINE

Great differences are found in the size of the large intestine of different animals. In the carnivora it is short and narrow, and possesses little or no cæcum. In herbivora the large intestine is well developed, with sacculated walls, and the cæcum is very large. Man occupies a somewhat intermediate position between these two classes. The great development of the large intestine in vegetable feeders is due to the fact that all the nutritious material in their food is enclosed in cells surrounded by cellulose walls. The cellulose is dissolved either by the agency of bacteria or by means of cytases in the vegetable cells themselves. In the horse and rabbit, for instance, the chief part of the digestion of the cellulose occurs by bacterial action in the cæcum.

In carnivora the large intestine has unimportant functions to discharge in digestion and absorption. The products of digestion are practically entirely absorbed by the time that the food has arrived at the ileocolic valve. In man the importance of the large intestine will vary with the nature of his food. If, however, a large quantity of vegetable food be taken, a considerable amount of material may reach the large intestine unabsorbed. A small proportion of this may undergo absorption in the large intestine, while the rest will pass out with the fæces, increasing their bulk. About 90 per cent. of the pectin of the food is absorbed, after being broken down by bacterial action in the colon.‡

The large intestine in man is the part of the alimentary canal in which micro-organisms flourish. In the mouth these are numerous, including various species of micrococci, staphylococci, streptococci, aciduric bacilli, spirochætes and spirilla, etc., the stomach has a very restricted flora in health, since the strong acid reaction of the contents kills most of the organisms which enter it. The upper small intestine also contains few; the ileum has a rapidly increasing bacterial content, *e.g.*, enterococci, coliform bacilli, staphylococci and yeasts, as the ileo-cæcal valve is approached. So far, however, there are few anaerobic bacteria and putrefaction is slight until near to the large intestine. Bacterial flora here becomes much more complex and numerous and includes very many species of aerobic and anaerobic bacteria as well as yeasts and protozoa. The commonest bacterial species are coliform bacilli, enterococci, staphylococci, aciduric bacteria, spirochaetes, sporing aerobes, and both pathogenic and non-pathogenic species of sporing anaerobic organisms, such as *Cl. welchii* and *Cl. putrifica*.

Bacteria are said to make up from a quarter to one-third by weight of the fæces, but the greater part are dead organisms.

At birth the intestine contains very few organisms, but is rapidly colonised both from mouth above and anus below. Intestinal bacteria are probably not essential to life, but aid in digestion and may also affect the host's well-being by their activity in respect of both vitamin synthesis and vitamin destruction.

The study of the products of the activity of these bacteria, and of the protection of the body from these and from direct invasion by the organisms themselves, forms an interesting chapter in pathology.

The mucous membrane of the large intestine is lined by simple tubular glands containing many typical "goblet" cells. Investigation by the fistula method shows that, when empty, the secretion of the large intestine is extremely scanty and composed wholly of mucus, while when stimulated mechanically by introduction of a drainage tube, a slight watery secretion is further contributed.

---

* Elliott. *J. Physiol.*, 1904, **31**, 157.
† Hurst. *J. Physiol.*, 1913, **47**, 54.
‡ Werch and Ivy. *Amer. J. Dig. Dis.*, 1941, **8**, 101.

The secretion in the cat's colon is effected through the action of the *nervi erigentes.*[*] Reflex secretion can be obtained by stimulating the central end of one nervus erigens, the other nerve being intact. We may suppose such a secretion, often copious, to be a result of the presence of irritating substances in the fæces. Secretion is also caused by pilocarpine, inhibited by atropine, and reduced by sympathetic stimulation. The secretion has a pH of 8·4 and is rich in mucin, but almost devoid of any enzymes. In the resting state the colon reabsorbs the entire secretion. The secretion of mucus aids the passage of the fæces along the gut.

As an absorbing organ the large intestine of man is of little importance. From observations on cæcal fistulæ in man it has been calculated that about 500 c.c. of water pass the ileocolic valve in the twenty-four hours. Of this about 400 c.c., together with the water secreted in the large intestine itself, undergoes absorption in the large intestine. The absorption of food substances introduced into this part of the gut is much slower than that which takes place on introduction by the mouth. In some cases after the introduction of large enemata into the large intestine, a certain amount may escape backwards into the ileum and may there undergo absorption. The isolated large intestine of man is able to absorb only about 6 grammes of glucose per hour and about 80 c.c. of water. If egg albumin or caseinogen solutions be introduced by the rectum, no absorption can be detected after several hours. Feeding by nutrient enemata is thus of little value.

It was formerly held that the large intestine excreted a considerable fraction of the ingested calcium and magnesium salts, and also iron and heavy metals when these were administered. So far as calcium, magnesium and iron are concerned this belief has been shaken by the careful work of Nicolaysen[†] and of McCance and Widdowson,[‡] who found that when their salts were injected into the blood there was no increased output of the elements by the bowel. In earlier work allowance was not made for failure of absorption when these salts were given by mouth.

Such absorption of *iron* as takes place occurs in the duodenum and upper part of the jejunum. Only about 0·1 mg. daily appears in the urine ; all the rest of the normal intake of about 7–8 mg. daily, together with any iron compounds given by mouth, is unabsorbed and appears in the fæces, chiefly as sulphide.

Of the acid radicals, *phosphates* may pass out either with the urine or with the fæces, the exact path taken being determined by the relative amount of calcium and alkali metals present in the food. If there is an excess of calcium most of the phosphates will leave with the fæces, but this does not imply that calcium phosphate has been excreted.

The large intestine does seem to be a channel of excretion for the heavy metals, such as bismuth and mercury. If bismuth be administered subcutaneously, the fæces will be found to contain this substance, and the wall of the large intestine will be stained black from a deposit of sulphide of bismuth. This deposit stops short at the ileocolic valve.

Movements of the Large Intestine. By means of the occasional peristaltic contractions, accompanied by relaxation of the ileocolic sphincter, the contents of the small intestine are transferred into the colon. This happens most rapidly when the next meal is taken, but in any case not till about four and a half hours after the last one. When food is swallowed, or when chyme enters the duodenum, a peristaltic rush is initiated in the small intestine ; on the arrival of the wave at the ileocæcal valve, the sphincter opens and allows the passage of some of the contents of the ileum into the colon (" gastro-ileal " reflex). In man the contents are considerable in bulk, are semi-fluid, and quickly fill up the transverse as well as the ascending colon.

---

[*] Wright, Florey and Jennings. *Q. J. Exp. Physiol.*, 1938, **28,** 207.

[†] Nicolaysen. *Skand. Arch. Physiol.*, 1934, 69 Suppl. 1 ; *Bioch. J.*, 1937, **31,** 107.

[‡] McCance and Widdowson. *J. Physiol.*, 1938, **94,** 148 ; *Bioch. J.*, 1939, **33,** 523.

Innervation. The innervation of the large intestine is derived from both sympathetic and parasympathetic sources; there are some differences in the details of anatomical distribution in its different parts, and in different species of animals. In man, the *colon* in its proximal part receives fibres from the superior mesenteric plexus; these are partly sympathetic, from the lumbar roots, and partly parasympathetic from the vagus. In its distal parts the colon is supplied by sympathetic fibres, which reach it from the upper lumbar roots *viâ* the pelvic splanchnic branches, intermesenteric plexus and nerves, the inferior mesenteric plexus, in which are many ganglia, and the inferior mesenteric nerves. The parasympathetic supply to the distal colon is from the second to fourth sacral spinal roots, and then by way of the hypogastric plexus, in which are also scattered ganglia, to the wall of the colon (Fig. 577).

Regarding the action of these nerves, there are discrepancies in the results of different investigators. It has often been claimed that the sympathetic is inhibitory to the musculature of the colon, with the exception of the ileocolic sphincter, to which it is motor, while the parasympathetic is motor to all except the ileocolic and anal sphincters. But there is doubt about the action of the sympathetic, which is without doubt often motor, at all events to the circular coat. Whether the vagus is motor to the cæcum is also doubtful.

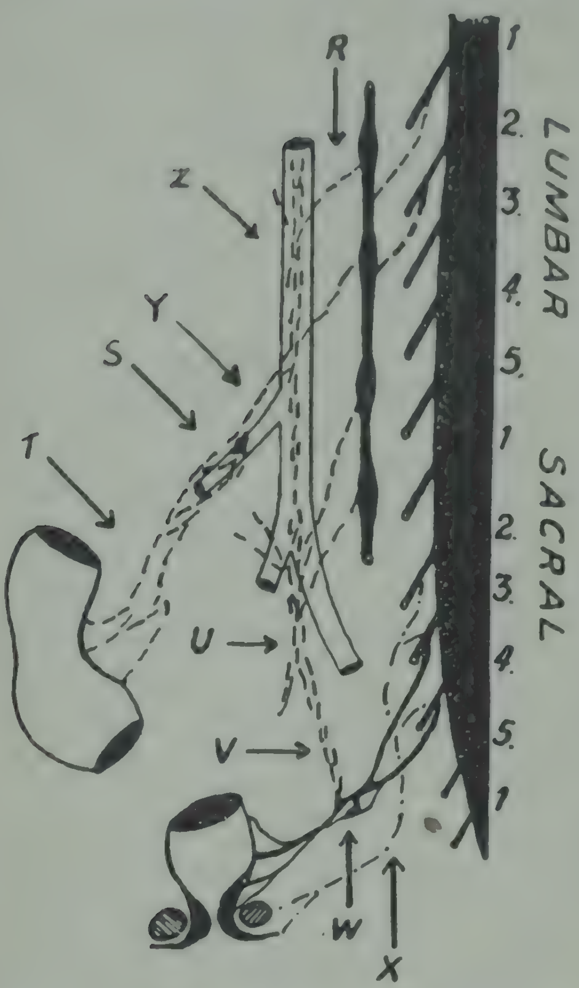


FIG. 577. The Nervous Outflow to the Distal Colon in Man. R., pelvic splanchnic nerves; S., inferior mesenteric plexus; T., inferior mesenteric nerves; U., presacral nerve; V., hypogastric nerve; W., hypogastric plexus; X., pudic nerve; Y., intermesenteric nerve; Z., intermesenteric plexus. (GARRY. *Physiol. Rev.*, 1934, **14**, 103.)

In addition to the nervous arrangements as described above, the large intestine has two intrinsic nerve plexuses, similar to those of the small intestine.

The rectum in man also receives a double nerve supply, of which the sympathetic is derived from the lumbar roots by way of the presacral nerve and the hypogastric nerves and plexus, while the parasympathetic supply from the sacral nerves also reaches it by way of the hypogastric plexus (Fig. 577). The *internal anal sphincter*, of plain muscle, has a similar supply, while the striated *external anal sphincter* receives a motor supply from the pudic nerve, which also supplies various other striated muscles of the perineum. Afferent fibres are present in the sacral and pudic nerves.

Movements. The movements of the large intestine may be observed by the X-ray method assisted, in man, by auscultation and percussion. Until about four hours after breakfast the cæcum is filled with gas, is resonant on percussion, but emits no sounds. At about the fourth hour the ileal contents begin to enter, producing characteristic sounds, and the X-ray shadow shows rapid filling of the cæcum and ascending colon; these show little peristaltic action but seem to be passively filled by the peristaltic contractions of the ileum.

The movements of the large intestine differ considerably in various animals,* according to the nature of the food and the part played by this portion of the gut in the processes of absorption. Man takes an intermediate position between herbivora and carnivora as regards his large intestine.

---

* ELLIOTT. *J. Physiol.*, 1904, **31**, 157; **32**, 401; ELLIOTT and BARCLAY-SMITH. *J. Physiol.*, 1904, **31**, 272.

In the cat there is active peristalsis and antiperistalsis of the ascending colon as it fills. The effect of the antiperistaltic waves is to force the food into the cæcum, regurgitation into the ileum being prevented by the tonic contraction of the ileocolic sphincter. The movements are rendered still more effective by the sacculation of the walls. The distension of the cæcum caused by this anti-peristalsis excites occasionally a true co-ordinated peristaltic wave which, starting in the cæcum, drives the food before it into the transverse colon. These waves die away before they reach the end of the colon, and the food is driven back again by waves of anti-peristalsis. The result of this process is the absorption of the greater part of the water of the intestinal contents, as well as of any nutrient material; and the drier part of the intestinal mass collects towards the splenic flexure, where it may be separated by transverse waves of constriction from the more fluid parts which are being driven to and fro in the proximal and intermediate portions. By means of occasional peristaltic waves these hard masses are driven into the distal part of the colon.

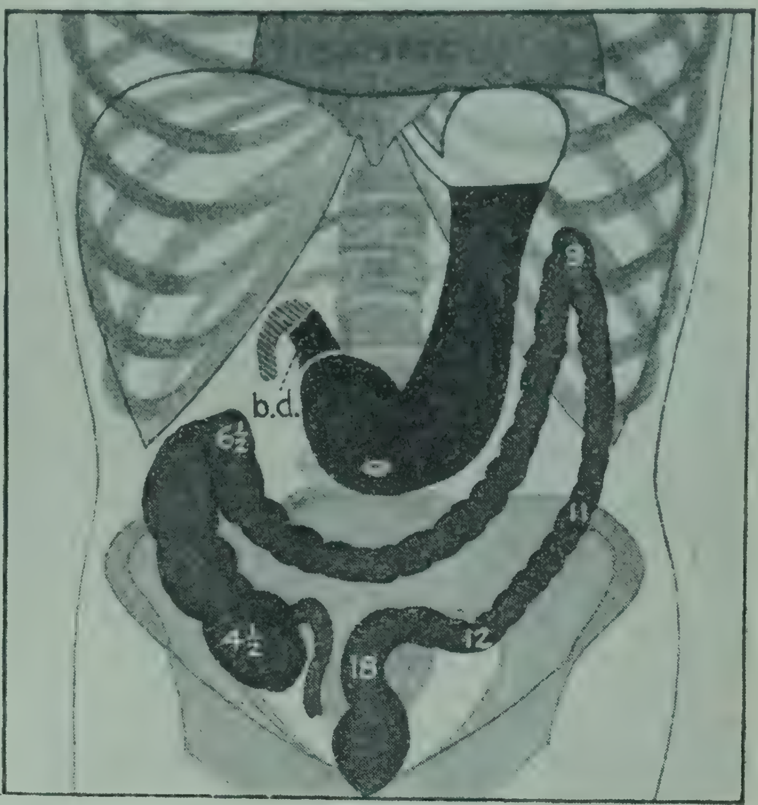


Fig. 578. Skiagram to show Normal Position of Colon in Man, and the Position attained by its Contents at different periods after a Meal containing Barium Sulphate.

The times of arrival at different levels are marked in hours after the meal on the colon. (Hurst.)

About 500 grammes of fluid chyme per day pass the ileocæcal valve in man; water and a little nutrient material is absorbed, chiefly in the cæcum and ascending colon, though the contents remain soft till they reach the pelvic colon, when they undergo inspissation, so that the average weight of fæces per day is about 135 grammes.

The contents, in man, reach the hepatic flexure on an average in about six and a half hours after the meal, the splenic flexure at about nine hours, and the pelvic colon in twelve hours, though there are great individual variations. The time taken for a barium meal to reach various levels in the colon in man is shown in the accompanying figure (Fig. 578). It should be remembered, however, that a barium meal is usually given after previously emptying the whole alimentary canal as completely as possible. In the normal course of events food residues are much slower in being completely eliminated in the fæces than these figures would indicate. Thus Alvarez and Freedlander, administering coloured glass beads with the food to normal adults, found that, even after four days, only about three-quarters of the beads had been recovered.

The progress of the contents along the large intestine is not a slow regular one, but occurs as a result of a few rather sudden, powerful contractions. These, due to what is called a gastro-colic reflex, occur especially when food is taken, and during defæcation, and are called "*mass peristalsis.*" By this means the contents of the cæcum and ascending colon may be moved bodily into the transverse colon in a few minutes, and similarly the contents of the transverse colon may be equally rapidly passed on into the pelvic colon and rectum.

Defæcation. Under normal conditions, the rectum and distal colon are subordinated to the spinal and other nerve centres, the inhibitory effect of the sympathetic supply being predominant, and leading to retention of the contents. Hence spinal anæsthesia leads to some increase of tone, whereas spinal transection in the lower thoracic region leads to complete relaxation. It would seem that the sympathetic supply has its centre in the lumbar cord, while the parasympathetic centre is higher up.

The normal stimulus for defæcation is probably the attainment of a certain distension of the rectum, which is reinforced by stimuli arising from the movement of fæces along the rectum and anal canal. These reflexly excite contraction of the colon, and relaxation of the sphincters, the former being due to the transmission of impulses along the sacral nerves. Stimulation of these nerves in an animal produces a rapid shortening of the distal part of the colon, due to contraction of the recto-coccygeus and longitudinal fibres of the gut, followed after some seconds by a contraction of the circular coat. This originates certainly as high as the splenic flexure, and possibly even higher, and spreading rapidly downwards empties the whole of this segment of the gut. In man, the pelvic colon becomes filled with fæces from below upwards, the rectum remaining empty till just before defæcation. In individuals whose bowels are opened regularly, the entry of fæces into the rectum gives rise to a sensation of fulness and acts as the call to defæcation. "Mass contractions" of the colon may occur as a reflex result of taking food into the stomach. The rectum thus becomes rapidly filled, and its distension produces the call to defæcation. Hence the act of defæcation after breakfast, or, in patients with irritable intestines, after any meal. If no response be made, the desire to defæcate passes away, since the rectum relaxes again. The minimal pressure required to produce the call to defæcate varies from 30 to 40 mm. Hg. The emptying of the rectum itself is largely assisted by the contractions of the voluntary muscles of the abdominal walls and pelvic floor. The last section of the rectum is emptied at the close of the act by a forcible contraction of the levator ani and the other perinæal muscles, and this contraction also serves to restore the everted mucous membrane.

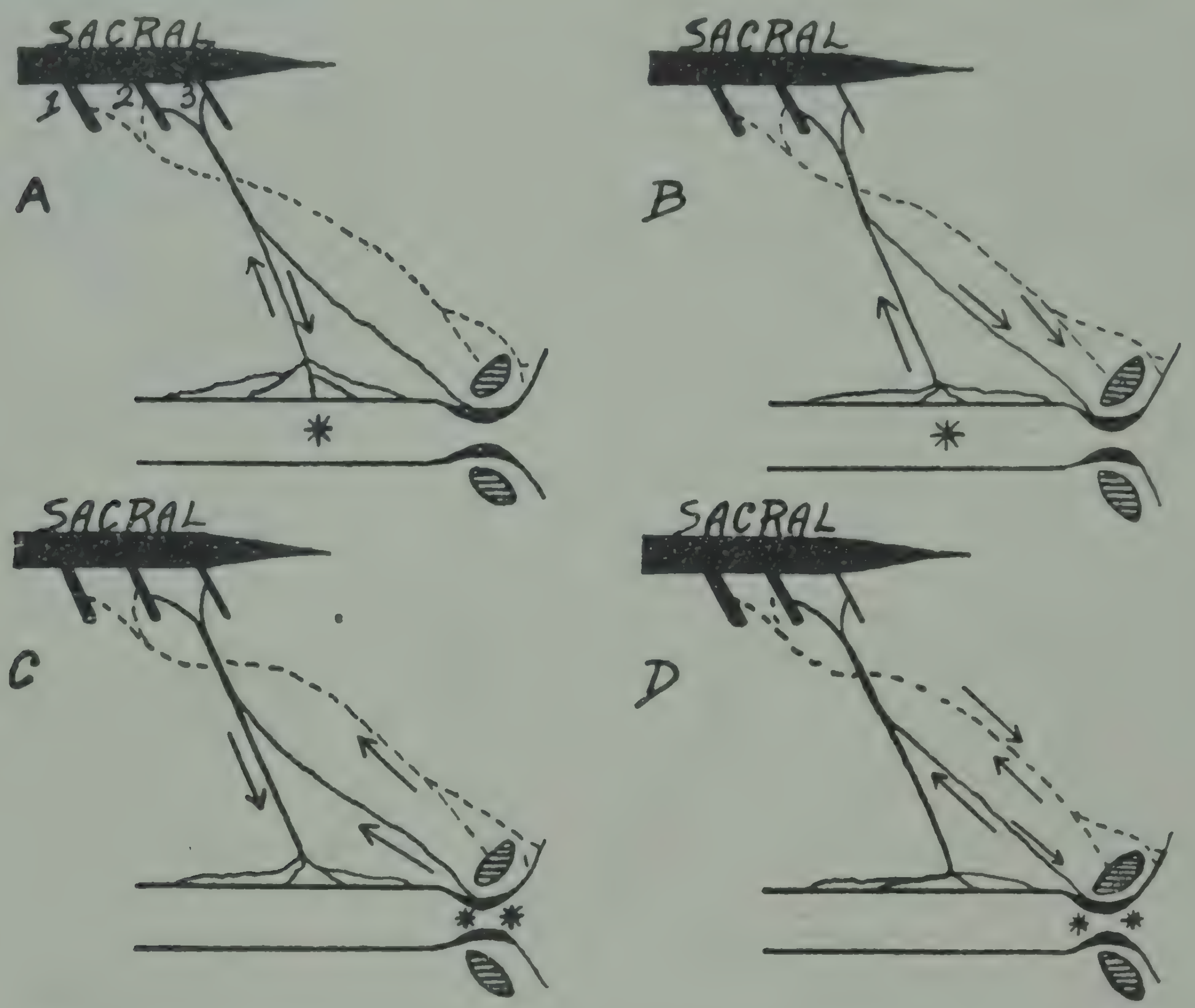


FIG. 579. Course of Impulses in Defæcation, according to Garry.
— pelvic nerve.
- - - pudic nerve.
* point of stimulation.

Stimulus to mucosa of distal colon causes (A) contraction of distal colon and (B) relaxation of both anal sphincters.

Stimulus to mucosa of anus causes (C) contraction of the distal colon and (D) relaxation of both sphincters.

The anal orifice is guarded by an internal anal sphincter of smooth muscle, and by a series of three striated external sphincters.* The internal sphincter is supplied by the sacral and hypogastric nerves, and the external anal sphincters by the pudic nerve. These sphincters are normally kept closed, even on section of the upper lumbar spinal cord. Stimuli in the colon or rectum which give rise to defæcation cause a reciprocal contraction of the lower bowel and relaxation of these sphincters.* Conversely, stimulation of the anal canal, as by movement of fæces through it, leads to reflex dilatation of the canal, and contraction of the rectum. As shown in Fig. 579, the afferent nerves in these reflexes are the pelvic nerves and the pudic nerves, the former being also motor to the colon and inhibitory to the internal sphincter, while the latter is motor to the external anal sphincter.

---

* MILLIGAN and MORGAN. *Lancet*, 1934, **227**, 1150.

The carrying out of this reflex act is normally dependent on the integrity of a certain part of the lumbar spinal cord. If this " centre " be destroyed, the tonic contraction of the sphincter muscles disappears. The centre may be either excited to increased action or inhibited by peripheral stimulation of various nerves, or by emotions such as fear. Application of warmth to the region of the anus causes reflex relaxation of the sphincter ; application of cold increases its tonic contraction.

Some time after destruction of the lumbar spinal centres a certain amount of automatic control is restored over defæcation. This has been supposed to be due to a fundamental control exerted by the local nerve plexuses,† but has not been finally explained.

The Fæces

In man or the carnivora the absorption of the constituents of a meal is practically complete by the time that the food has arrived at the lower end of the ileum. The fæces in the main are not derived from the food, but are produced very largely in the alimentary canal itself. This is shown by the fact that on analysing the fæces no soluble carbohydrates or proteins, albumoses, peptones or amino-acids are to be found. After a meal of meat, microscopic examination of the fæces reveals no trace of striated muscle fibres. Moreover, animals in a state of complete starvation form fæces which do not differ in their composition from the fæces which are found after feeding, though the amount is less. In one experiment Hermann isolated a loop of gut, joining its ends together so that a continuous ring was formed. The continuity of the gut was then restored by suturing the two free ends. After some weeks the isolated loop was found to contain a semi-solid material similar to fæces.

So long as vegetables or coarsely ground cereals are excluded from the diet, the nature of the latter does not alter the chemical constitution or appearance of the fæces. Under these circumstances the fæces have the following composition :

Water	65 to 67 per cent.
Nitrogen	5 to 9 ,,
Ether extract	12 to 18 ,,
Ash	11 to 22 ,,

The ash consists chiefly of lime and phosphoric acid with some iron and magnesia. The ethereal extract contains fatty acids and a small amount of lecithin. Neutral fat is present in very small proportions. The fæces also contain small quantities of cholic acid and its decomposition product, dyslysin, also coprosterol, a reduction product of cholesterol, and a certain amount of purine bases, consisting of guanine, adenine, xanthine and hypoxanthine. On the average the fæces contain about 0·11 gramme of purine bases per idem, about seven times as much as is contained in the urine passed in the same time. The material basis of the fæces seems to be largely inspissated mucus, bile and other secretions, desquamated epithelial cells from the intestinal wall, and bacteria, of which countless numbers, chiefly dead, are present. It has been reckoned that as much as 50 per cent. by weight of the fæces may consist of dead bacteria.

---

* Garry. *J. Physiol.*, 1932, **74,** 14P ; *Physiol. Rev.*, 1934, **14,** 103.
† Denny-Brown and Robertson. *Brain*, 1935, **58,** 256.

Very different is the composition of fæces if the food contains a large amount of cellulose. Not only does the ingested cellulose pass unchanged into the fæces, but large quantities of other substances enclosed in the cellulose walls may also escape digestion. Moreover, the increased bulk of the undigested residue stimulates peristalsis, and thus quickens the passage of the food through the gut. The influence of the character of the food is well illustrated by a comparison of the amount and composition of the fæces on different kinds of bread (RUBNER) :

Kind of bread.	Weight of moist fæces.	Weight of dried fæces.	Percentage of ingested food.	Nitrogen (g.).
Bread from fine flour .	132·7	24·8	4·03	2·17
Bread from coarse flour .	252·8	40·8	6·66	3·24
Brown bread . . .	317·8	75·79	12·23	3·80

The indigestible cellulose in the food is not without value. If the food is capable of entire digestion and absorption, the amount of fæces formed is limited to that produced by the intestinal wall itself. The small bulk exercises very little stimulating effect on the intestine, and the movements of the latter will therefore tend to be sluggish. The presence of a certain amount of cellulose in the diet may therefore be of considerable advantage by giving bulk to the fæces and ensuring the proper regular evacuation of the lower gut.

Intestinal Gas. Gas gets into the intestine from swallowed air, by diffusion from the blood, or as a result of fermentative changes. The longer it stays there the more nearly it gets to equilibrium with the blood. The total quantity of gas passed daily *per anum* is very variable ; a normal average is about 500 c.c.*

10. The Sensations of Hunger and Thirst. When the body is in need of food or water, the characteristic sensations of hunger or thirst arise and increase in intensity if the need is not satisfied. It is important to distinguish between hunger proper and appetite : the former is unpleasant or even painful, while the latter, though accentuated by actual hunger, is mainly psychological and generally pleasurable.

It is not certain what is the fundamental cause of the hunger sensations. Hunger may be specific, as well as general, as, for example, the characteristic salt-hunger which leads animals in the wild state to travel long distances in search of its relief. Indeed, the constant average dietary balance in man and in animals of all species must in the long run be due, not so much to any conscious selection of the various constituents as to an unconscious choice guided by vague specific desires, now for more carbohydrate, now for salt, or fat, or protein. We have no knowledge of the mechanisms upon which these specific hungers, or their satisfaction, depend, nor could it be denied with any certainty that general hunger is due to a summation of such specific hungers, though probably it is not. The view that general hunger is associated with a lowering of the blood sugar receives some support from the fact that the blood sugar is a common pabulum into which certain other nutritional materials are convertible, so that its concentration in the blood may well serve as an index of the nutritional needs of the body. In support of this view is the fact that a sensation of hunger is one of the first effects of lowered blood sugar resulting from the action of a dose of insulin, though against it is the fact that hunger is a very prominent symptom in diabetics in whom, however, the blood sugar is permanently raised.

---

* BEAZELL and IVY. *Journ. Dig. Dis.*, 1941, **8,** 128.

As regards the actual pangs of hunger, which are referred to the epigastric region, there is little doubt that these are caused by powerful contractions of the body of the stomach, which begin as soon as the organ is empty or nearly empty,* at first as feeble contractions at intervals of twenty to thirty seconds, but finally becoming very powerful and lasting for about thirty seconds. Each contraction is accompanied by a hunger pang, and by an increased secretion of saliva. Vagotomy and sympathectomy in man each abolish hunger pangs, the former by depressing gastric motility, the latter by interrupting the afferent path ; but vagotomy certainly does not abolish sensations of hunger due to insulin.†

The effects of deprivation of water are evident sooner than those of want of solid food, and are correspondingly more urgent. In the first stage there is craving for liquid, and a feeling of dryness in the mouth ; later the dryness increases, the air breathed seems hot and dry, and frequent and painful swallowing movements are made ; food cannot be swallowed and even sugar may not dissolve in the mouth. Delirium and death soon follow, the total duration of life without water depending on the rate of loss and so being brief (one to two days) after exertion in hot dry climates, but longer, up to two weeks or more, in cooler damper climates and under conditions of rest.

The thirst of *diabetes mellitus* and of *diabetes insipidus* is imperative : in the former case the water is lost in company with the sugar excreted; in the latter there is a primary disturbance of the normal water distribution of the body, so that water which should normally be stored, *e.g.* in the connective tissues and liver, is excreted and leads to a constant shortage.

It has often been supposed that thirst is a general sensation arising in some way from the desiccated tissues. This view received support from experiments by Claude Bernard, who found that in animals with double œsophagostomy, as for sham feeding, or animals with open gastric fistulæ through which fluid escaped on swallowing, thirst was apparently not relieved by drinking. It has certainly been found that after long deprivation of water the osmotic pressure of the blood is demonstrably increased,‡ and possibly this is the case in diabetes mellitus, yet it is not possible to show any clear change in that direction in the early stages, in which, nevertheless, thirst is well established. Probably, therefore, the changes in the blood are too small to detect by physico-chemical methods, though sufficient to produce the characteristic sensations of thirst.

The alternative view, advanced by Cannon,§ is that the sensation of thirst is, as it appears to be, due to a dryness of the mouth, from any cause whatever. Owing to the passage of the inspired air through the pharynx, this always tends to become dried, but drying is ordinarily prevented by the secretion of saliva. When the tissues are deprived of water, however, the secretion of saliva is reduced, thus leading to the slight drying of the pharynx and producing the sensation of thirst. But we do not know how this result is produced. For example, in men, intravenous injection of 300 c.c. of 5 per cent. NaCl causes thirst, and reduction of salivary flow ; but analyses showed an increase of 5–10 per cent. in plasma and extracellular volume.‖ The thirst and dry mouth were prevented if 500 c.c. water was given half an hour before the salt injection, but the plasma changes were similar to those not preceded by water drinking.

Administration of atropine, by checking the secretion of saliva, gives rise to thirst, and conversely pilocarpine when administered to thirsting animals relieves it (Pack).¶ It has been stated also that after ligature of the parotid ducts, dogs show constant thirst. But doubt has been cast on this by Montgomery,* who finds that removal of all the salivary glands from dogs does not alter the water intake. The same is true of men with no salivary glands ; they drink no more than others, but drink more often.†

---

* Carlson. "Control of Hunger in Health and Disease." Chicago, 1916.

† Grossman and Stein. *J. Appl. Physiol.*, 1948, **1**, 263.

‡ Mayer. *C. R. Soc. Biol.*, 1900, **52**, 153.

§ Cannon. *Proc. Roy. Soc.*, 1918, **90B**, 283.

‖ Holmes and Gregerson. *Amer. J. Physiol.*, 1947, **151**, 252.

¶ Pack. *Amer. J. Physiol.*, 1923, **65**, 346.

For Reference

ALVAREZ (1940). An Introduction to Gastro-Enterology. (London : Heinemann.)
ALVAREZ (1924). Intestinal Autointoxication. *Physiol. Rev.*, **4**, 352.
BABKIN (1944). The Secretory Mechanism of the Digestive Glands. (London : Hoeber.)
BABKIN (1928). The Digestive Work of the Stomach. *Physiol. Rev.*, **8**, 365.
BARCLAY (1936). The Digestive Tract. (Cambr. Univ. Press.)
CARLSON (1916). The Control of Hunger in Health and Disease. (Univ. Chicago Press.)
GARRY (1934). The Movements of the Large Intestine. *Physiol. Rev.*, **14**, 103.
IVY (1930). The Rôle of Hormones in Digestion. *Physiol. Rev.*, **10**, 282.
MARKOWITZ (1937). Text-book of Experimental Surgery. (Baillière, Tindall and Cox.)
PAVLOV (1902). The Work of the Digestive Glands.
ROGER and BINET (1931). Traité de Physiologie Normale et Pathologique. Tome II. (Masson, Paris.)
RYLE (1926). Gastric Function in Health and Disease. (Oxford Med. Pub.)
VAN LIERE (1941). The Effect of Anoxia on the Alimentary Tract. *Physiol. Rev.* **21**, 307.

---

* MONTGOMERY. *Amer. J. Physiol.*, 1931, **96**, 221.
† STEGGAERDA. *Amer. J. Physiol.*, 1939, **126**, 635.

CHAPTER XLV

THE ABSORPTION OF THE FOODSTUFFS

The processes of digestion result in the conversion of the insoluble constituents of the food into soluble substances of smaller molecular weight, which can pass into the blood stream, and so become available for use in the body. The transfer of the dissolved materials, together with water, into the blood and body fluids, is called absorption.

Absorption from the Stomach. The rapidity with which highly diffusible substances, such as ethyl alcohol or hydrocyanic acid, produce their physiological effects, even when the pylorus has been obstructed, shows that some absorption can occur from the stomach. The substances named, however, have definite vapour pressures, and might pass as vapours.

If we may judge by experiments with heavy water (D_2O), there may be quite measurable absorption of water from dog's stomach pouches.*

The chief absorption of water occurs in the small intestine. It is on this account that the salient symptoms of dilatation of the stomach with stenosis of the pyloric orifice can be nearly all referred to the deprivation of the body of water, and can be often relieved by the administration of water, either subcutaneously or by the rectum, *i.e.* by the channels through which absorption is still possible.

Sugars and amino-acids may perhaps also be slowly absorbed in the stomach.

Absorption from the Small Intestine. By far the greater part of the absorption of all materials takes place in the small intestine. The inner surface of the small intestine is greatly augmented by the formation of folds, the *valvulæ conniventes*, and by the protrusion of villi from the surface. It has been calculated that each square millimetre of intestine represents an absorbing surface of 3 to 18 mm.^{2}, and a total capillary surface of about the same order, though not all these capillary surfaces are superficial in position.

Structure of Villi. Each villus (Fig. 580) consists of a framework of reticular tissue containing many leucocytes in its meshes, separated from the lumen of the gut by a continuous layer of columnar cells. These surface cells, which rest on an incomplete basement membrane, are of four kinds. Much the most abundant are the *principal* or *columnar absorbing cells*. These present a striated border or fringe on their outer sides (Fig. 580). Scattered among them are numerous *goblet cells*, *Paneth cells* and *enterochrome* or *basal granular cells*.† Similar groups of cells are found lining the crypts of Lieberkühn. Apart from the fact that the goblet cells secrete mucus, the function of these various cells is a matter for speculation.

The villus offers two channels by means of which material which has passed through its epithelium may be carried away. In the middle of the villus is the central lacteal, which leads into a plexus of lymphatics placed superficially to the *muscularis mucosæ*. From the superficial plexus communicating branches pass vertically to a corresponding plexus lying in the submucosa. Valves are present in abundance in the deeper plexus, so that fluid can pass easily from the lacteal to the deeper plexus, but not in the reverse direction. From the muscularis mucosæ, unstriated muscle fibres pass up through the villus, to be attached partly to the outer surface of the central lacteal, partly by expanded extremities to the basement membrane covering the surface of the villus. Contraction of these muscle fibres might empty the central lacteal into the deep plexus of lymphatics and may also cause an expulsion of the contents of the spaces of the retiform tissue of the villus into the central lacteal. Formation of lymph is constantly proceeding, and its flow is doubtless aided by the continual movements of the muscular walls of the canal.

---

* Cope *et al.* *J. Clin. Inv.*, 1943, **103**, 22.

† Macklin and Macklin. In "Cowdray's Special Cytology," 1935, p. 283.

The other possible channel of absorption is by the capillary blood vessels of the villus. Each villus is supplied with blood from one or two arterioles which break up into a rich plexus of capillaries lying close under the basement membrane of the villus. The total surface area of the portions of the outer capillary walls which lie immediately below the epithelium is about one-third of the whole epithelial surface, so that the surface through which interchange may occur between the contents of the bowel and the blood is at least three to five times the area of the intestinal mucosa in plan. In these capillaries the blood is circulating rapidly, so that a considerable amount of material may pass into them without appreciably altering the composition of the blood. But the blood in these vessels is at a high pressure, probably not less than 30 mm. Hg. so that we shall have to place out of court any possibility of absorption occurring by a process of filtration. The return blood is collected into one or two veins, which run to a submucous plexus and then join the radicles of the portal vein in the mesentery. It will be evident that, so far as processes of diffusion are concerned, the passage of substances from the lumen of the intestine into the blood stream would be much more favoured than passage into the lymph, for the blood current is so much more swift, and its surface of exposure so much greater than that of the lymph.

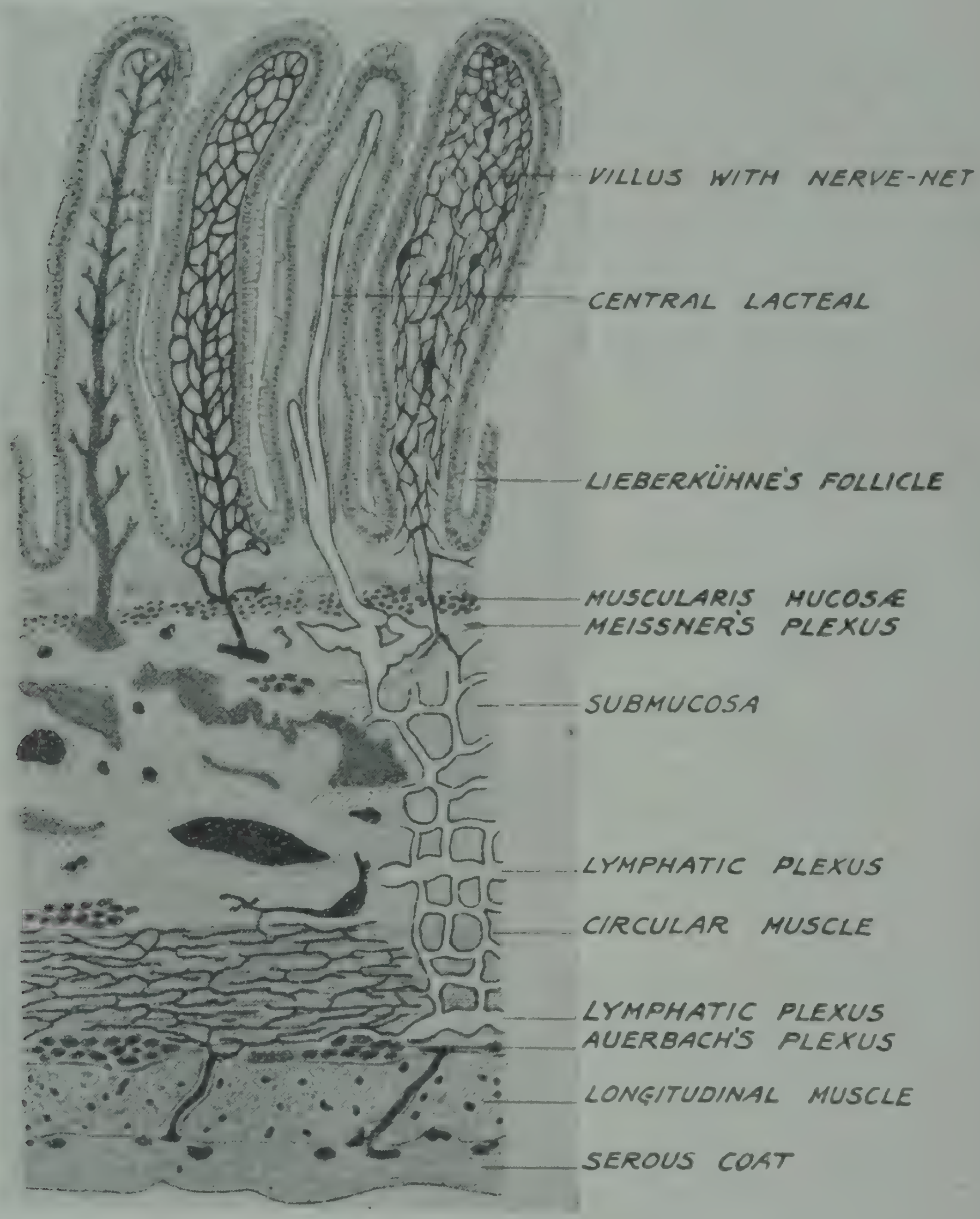


FIG. 580. Diagrammatic Section through Wall of small Intestine to show Vascular and Lymphatic Arrangements. (After SCHAFFER.) (*Lehrb. d. Histol. u. Histogenese*, 1932.)

Movements of Villi. One of the most striking facts in connection with intestinal absorption is that when it is in progress the villi are in a state of continual movement,* which, in a chloralosed dog, can easily be seen on opening the intestine opposite the mesenteric border and examining the mucosa by reflected light (Fig. 581). The movement in carnivora consists of a shortening, at rather irregular intervals of time (about six per minute) to about half the extended length, and then a more gradual extension; the cause of the shortening is no doubt the contraction of the plain muscle of the villus, but the lengthening has not been explained. The villi move quite independently of one another. In herbivorous and omnivorous animals the movements are somewhat different.

In the fasting state the villi are motionless and flattened, so as to look elliptical when seen from above, but if moistened with water or Ringer's solution they immediately swell out to a rounded cross-section and begin their pumping movements. The movements usually cease soon after removal of the intestine from the body, after loss of blood, stoppage of circulation, asphyxia or acapnia.

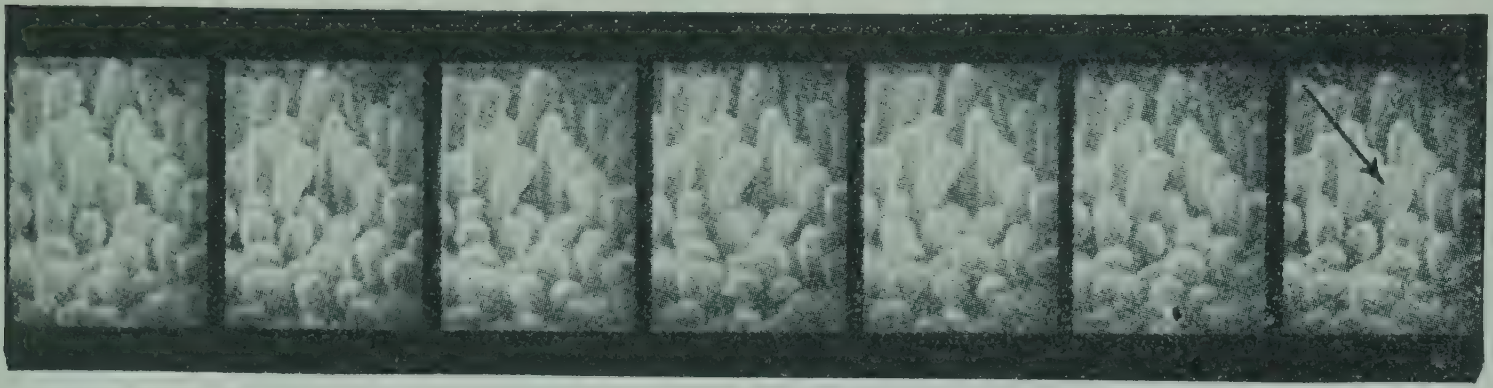


FIG. 581. Portion of cinematomicrograph film showing movements of villi in living intestine of dog during somewhat less than one second. Note the villus indicated by the arrow. (KOKAS and LUDÁNY, *Pflüger's Arch.*, 1930, **225**, 421.)

Mechanical stimuli to the base of the villus cause a contraction, probably by exciting the Meissner's plexus, which appears to be motor to the plain muscle of the villi. Chemical stimuli, *e.g.* eserine, yeast extract, amino-acids, etc., also cause rhythmic contractions of the villi. When the mucosa is stripped off from the submucosa these stimuli all fail if the muscularis mucosæ, together with Meissner's plexus, is detached, but remain effective when it is left behind. The natural stimulus for the contractions of the villi is stated to be a hormone, called *villikinin.* It is extractable from the intestinal mucosa by dilute hydrochloric acid and has been shown, by cross-circulation experiments, to be present in the blood after hydrochloric acid has been placed in the duodenum of a donor dog.†

The Absorption of Water and Dissolved Substances. The absorption of water from the alimentary canal is very little influenced by the water content of the body, the greater part of the ingested fluid being absorbed in the intestine. Thus a man may maintain the water balance of his body steady, whether he take one litre or six litres of water daily. The regulation, apart from that determined by thirst, is carried out by the kidneys.

The intestinal contents at the ileocæcal valve contain relatively nearly as much water as they do at the upper part of the jejunum. Their absolute bulk, however, is much smaller, so that only a small proportion of the water that has been taken in by the mouth remains to be absorbed in the large gut—an amount considerably less than that which has been added to the contents of the small intestine in the form of secretions.

When salt solutions are introduced into the small intestine, they are rapidly absorbed without any increase in the rate of lymph flow, whereas the absorption of large amounts of fluid may cause dilution of the plasma. This has usually been interpreted as meaning that the blood vessels at the surface of the villi are solely responsible for the absorption. But an alternative explanation is possible.

---

* HAMBLETON. *Amer. J. Physiol.*, 1914, **34**, 446; KING and ARNOLD, *Amer. J. Physiol.*, 1922, **59**, 97; KING, ARNOLD and CHURCH, *ibid.*, **61**, 80; KOKAS, *Pflüger's Arch.*, 1930, **225**, 416, 421; 1932, **229**, 486; VERZÁR. *Ergebn. d. Physiol.*, 1931, **32**, 391; KOKAS and LUDÁNY, *Pflüger's Arch.*, 1932, **231**, 20 (isolated intestine).

† KOKAS and LUDÁNY. *Q. J. Exp. Physiol.*, 1938, **28**, 15.

On its way through the sub-mucosa the lymph vessel comes into very close contact with special plexuses of small veins, so that it would be quite possible for transfer of water or dissolved material of the lymph to take place into the blood. Hence, in addition to absorption from the intestinal contents into the blood, there might also be reabsorption from lymph into blood.

The problem as to the forces which are responsible for the transfer of various substances from the lumen of the intestine across the " membrane " consisting of epithelial cells, together with interstitial tissue and endothelium of blood or lymph capillary, into blood or lymph respectively, is a very old one to physiology and still presents many difficulties. As regards the *rate* at which the transport occurs, it must be granted that this will be largely influenced by the contractile activity of the villi referred to above. Setting this aside, however, as likely to affect only the quantitative aspect, the problem still allows of no facile solution by reference to simple laws of physical chemistry, probably because our understanding of those laws is not yet sufficiently complete. Opinions, therefore, run in two directions ; according to one view the living intestinal membrane selects those constituents which are to be absorbed in accordance with the needs of the organism ; the other school considers that though the process of absorption may be vastly complicated by special physico-chemical properties of the membrane, yet it is not at present desirable to regard it as fundamentally different from that concerned in the passage of substances through dead membranes. Some of the factors concerned in transport through membranes in general have already been discussed in Chapter VII.

The special physico-chemical phenomena which need some further comment here may be grouped under the headings : (1) Diffusion and Osmosis ; (2) The Special Properties of the Membrane ; (3) Hydrotropy.

(1) Diffusion and Osmosis. These terms have already been explained in Chapter VII., and the reader is here only reminded of some pertinent facts which are in apparent contradiction to the accepted laws of osmosis and diffusion. As an example we may take some experiments * on the passage of water and sodium chloride from the intestine of dogs with Thiry-Vella fistulæ, during half-hour periods, after placing 75 c.c. of solution in the loop. When distilled water was placed in the gut, only 59 per cent. of it was absorbed, but when 0·4 per cent. NaCl was used, as much as 95 per cent. of the water was absorbed. With ascending strengths of NaCl solution, the water absorbed rose from 59 per cent. with pure water to a maximum of 95 per cent. between 0·4 and 0·7 per cent. NaCl, and then fell rapidly because with concentrations over 1·8 per cent. osmotic passage of fluid *into* the intestine, as well as considerable secretion, took place. The absolute amount of sodium chloride absorbed increased with the concentration of NaCl placed in the intestine, up to 1·5 per cent., *beyond which it fell* ; but the percentage of the NaCl given which was removed rose with concentration until about 0·7 per cent. and then fell. At 0·1—1·2 per cent. NaCl (when allowance for secretion into the gut was made) the fluid was absorbed without change of composition, *i.e.* water and NaCl were absorbed together ; from solutions over 1·2 per cent. NaCl, however, the chloride was absorbed faster than the water. Some of these facts might be explicable on the simple laws of osmosis, but some certainly cannot be so explained. The absorption of isotonic solutions has been supposed to be accounted for by the osmotic pressure of the colloids in the blood plasma, but the rate of absorption of 0·9 per cent. NaCl solution was not altered by the addition to it of 6 per cent. gum acacia, which provides a colloid osmotic pressure about equal to that of the plasma.

When isotonic solutions of sodium chloride and sodium sulphate are placed in an intestinal loop the sodium chloride is almost completely absorbed, although blood plasma contains at least 0·6 per cent. of that salt, whereas the sodium sulphate is unabsorbed and remains in approximately isotonic solution.

---

* Rabinovitch. *Amer. J. Physiol.*, 1927, **82**, 279.

Similarly, magnesium ions are unabsorbed.* It is clear, therefore, that osmotic work is performed by the intestine even in an apparently simple case like this one. The active intervention of the cells in the absorption of any substance can be abolished by any means which destroys their vitality, such as the addition of sodium fluoride or cyanide to the fluid to be absorbed, or destruction of the epithelium by previous temporary occlusion of the blood vessels supplying the loop of intestine: when poisoned by cyanide, the difference between the intestinal mucosa and a dead membrane disappear.

Wallace and Cushny,† as the result of their experiments on the relative absorbability of various sodium salt solutions from the gut, divide the salts into four main classes as follows: (1) Sodium chloride, bromide, iodide, formate, acetate, propionate, butyrate, valerianate, caprate; (2) Ethyl sulphate, nitrate, lactate, salicylate, phthalate; (3) Sulphate, phosphate, ferrocyanide, caprylate, malonate, succinate, malate, citrate, tartrate; (4) Oxalate, fluoride.

Of these the first class contains salts which are absorbed with great ease from the intestine. The second are absorbed with somewhat greater difficulty. The third group are absorbed so slowly, *i.e.* the salts retain the water in which they are dissolved so long, that they act as laxatives. The members of the fourth class are not absorbed at all. It is evident that this classification is independent of the diffusibility of the salts. There is, however, one character which apparently determines the non-absorbability (relative or absolute) of the members of the third and fourth classes. All these salts form insoluble compounds with calcium. This common character is not an explanation of the permeability of the cell wall, but is simply a general statement of one of the conditions which affect the power of the cells to take up salts from their solutions, this power being absent in the case of salts which furnish an insoluble calcium compound.

When sugar solutions are used, both sugar and water are completely absorbed, but again certain anomalies appear, which are not readily explained on physico-chemical grounds. Glucose disappears almost as rapidly as sodium chloride or iodide, though its physical diffusibility is less than that of these salts. Further, great differences, which are also not referable to the diffusibility of the sugars in question, are found between the rates at which different sugars are absorbed and in different species of animals.‡ In cats the order of speeds of absorption is galactose $>$ xylose $=$ glucose $>$ fructose. In rats, galactose is the most rapidly absorbed, then comes glucose, while fructose is only half, and mannose only one-fifth as rapidly absorbed as glucose,§ the actual ratios of rates being 110, 100, 43 and 19 respectively. The pentoses xylose and arabinose are very slowly absorbed (15 and 9 respectively); when the intestine is dead, and used like a dialyser, the rates are different, for then, though all substances pass more quickly, xylose passes out even more rapidly than glucose.‖ There is never any passage of glucose outwards from the blood into the lumen of the bowel.

When the rate of absorption of glucose from solutions of different strengths is studied, it is found that the most rapid absorption is from 0·75M (13·5 per cent.) solution, the rate of absorption from this strength being about 20 per cent. more rapid than from 2M solutions.¶ A still greater rate of absorption was obtained when 0·2 per cent. phosphate was added to the 0·75M glucose solution. There is considerable evidence that glucose, and probably other sugars, are absorbed in the form of combinations with phosphoric acid. Thus, serum phosphate is increased during absorption,** while if the phosphatase be poisoned, absorption of glucose from the intestine is inhibited.††

As regards the disaccharides, it would seem that unless these are hydrolysed by intestinal enzymes there is no absorption at all, though they are diffusible readily enough through dead membranes.

---

* Ingraham and Visscher. *Amer. J. Physiol.*, 1938, **121**, 771; Ingraham, Peters and Visscher. *J. Physical Chem.*, 1938, **42**, 141.

† Wallace and Cushny. *Amer. J. Physiol.*, 1898, **1**, 411; *Pflüger's Arch.*, 1899, **77**, 202.

‡ Davidson and Garrey. *J. Physiol.*, 1940, **97**, 509.

§ Cori. *Physiol. Rev.*, 1931, **11**, 145.

‖ Auchinachie, Macleod and Magee. *J. Physiol.*, 1930, **69**, 185.

¶ Magee and Reid. *Journ. Physiol.*, 1931, **73**, 163.

** Bodansky. *J. Biol. Chem.*, 1934, **104**, 473.

†† Lundsgaard. *Biochem. Ztschr.*, 1933, **264**, 221.

It is a further curious fact that some substances, *e.g.* glucose, or amino-acids, are absorbed at a constant rate from the intestine, in spite of the fact that as absorption proceeds the concentration of the substance in the gut diminishes, yet other substances, such as lactic acid and iodides, are absorbed at rates which depend on the concentration of the substance in the intestine at the time.*

When we turn from these facts to seek physico-chemical explanations in accordance with the laws of osmosis, it is necessary, if we are to make any further progress, to consider the latter in a wider sense. In the narrower sense, osmosis refers only to the passage of water through membranes impermeable to all solutes, but it is clear that we need to consider the passage of both water and solutes through membranes of restricted, but variable, permeability. These laws are complex, and have been specially studied by Schreinemakers,† who finds that the course of the interchange of water and solutes will vary, and in an unpredictable manner, according to the number of solutes, their concentration, and the nature of the membrane, so that it is quite possible, for instance, even with dead membranes, to get water or solute moving in a direction which is apparently opposed to simple laws of osmosis as deduced from semipermeable membranes. Experiments with radiosodium (Na^{24}) as a tracer ion show that sodium is constantly moving in and out of the gut, whatever the net direction may be, the total turnover rate between intestine and blood being about equal to the whole plasma sodium every eighty-three minutes.‡ It would be beyond our scope here to enter further into this aspect of the subject.

We must conclude that an active transference of water and solutes from the lumen into the blood stream is effected by the intermediation of forces having their origin partly in the metabolism of the cells themselves. This work of absorption by the cells may be aided or hindered according to the physical conditions present, though our knowledge of these physical conditions is at present too fragmentary to enable us to appraise the respective parts played, on the one hand by the physical conditions and on the other hand by the activity of the cells, which primarily results from and secondarily modifies them by introducing fresh ones. If these act against the cells, *e.g.* if the fluid be hypertonic, the absorption is effected more slowly, while with hypotonic solutions the physical conditions concur with the vital activity of the cells in bringing about a very rapid transference of fluid from the gut into the blood vessels. Among these physical conditions we must reckon the nature of the substances present in the intestine. If these can pass easily into, and through, the cells, *e.g.* ammonium, sodium, or chloride ions, absorption is carried out rapidly. If, on the other hand, the solutes in the intestinal contents are but slightly diffusible, or have very little power of penetrating into the cells, the absorption of water by the cells causes an increased concentration of the solutes, and therefore an increased osmotic pressure which offers a resistance to any further absorption.

(2) **Special Properties of the Membrane.** As already explained in Chapter VII., the results of osmotic experiments, even with dead membranes, depend very much on the properties of the membrane, and this in its turn is dependent, not only on its chemical composition, but, since the membranes with which we are concerned are essentially colloidal, also on its previous history.

Like other cells, those of the intestinal epithelium are bounded on their free surface by a " lipide " layer, permeable only by such substances as are soluble in lipides. On the other hand, the cement substance between the cells may be of a different character and possibly permeable to water-soluble substances. The question has been propounded whether the greater part of the substances which enter the blood plasma from the gut pass between the cells or through them. Water could probably pass in either way. Most of the inorganic salts, as well as the sugars, are insoluble in lipoids and would have to pass between the cells. When the question is investigated by the use of dyestuffs, soluble or insoluble in lipides, it is found that the lipide-soluble dyestuffs, such as neutral red or toluidin blue, pass into the cells, whereas the dyestuffs which are insoluble in such substances pass into the intercellular spaces. Too much stress, however, must not be laid on these experiments. The fact that absorption of these dyestuffs is determined by the physical conditions of the cell membrane is no proof that the absorption of the normal food constituents is determined in the same way. In fact, it is quite legitimate to assume that the lipide limiting layer round every cell has as its main office, not the regulation of the access of foodstuffs to the cell, but its protection from anything which it does not require. If it were not for such a membrane the assimilation of a salt would be determined entirely by its concentration in the immediate surroundings of the cell, whereas we know that assimilation by any living organism, whether uni- or multi-cellular, is regulated in the first place by the activity of the organism itself. The properties of the membranes are also liable to be altered by such phenomena as adsorption, electric charges, imbibition, Donnan equilibria, etc.

---

* Cori and Cori. *J. Biol. Chem.*, 1928, **81**, 389.

† Schreinemakers. *J. Gen. Physiol.*, 1928, **11**, 701 ; 1929, **12**, 555 ; 1930, **13**, 335.

‡ Visscher *et al.* *Amer. J. Physiol.*, 1944, **141**, 488.

(3) **Hydrotropy.** Hydrotropy* is the property of certain substances to render other substances soluble in water, in which otherwise they would be insoluble. The increase of solubility is sometimes remarkable, as, for instance, a 100-fold increase in the solubility of sulphonal by sodium salicylate, a five-fold increase for uric acid by atophan, etc. Hydrotropic substances are usually salts of certain organic acids which powerfully lower surface tension ; the substances which are brought into solution by their agency include fats, lipides and sterols.

Verzár and Kúthy† showed that various substances, particularly fatty acids, are dissolved, and, what is much more important, rendered dialysable, by the bile salts. Thus oleic acid was emulsified in a phosphate buffer-solution, and then 5 per cent. sodium glycocholate added. The mixture was placed in a dialyser, outside which was the same mixture without any fatty acid. After sixteen hours' dialysis the contents of the dialyser and the outer fluid were analysed for fat, with the results shown :—

DIFFUSION OF OLEIC ACID DISSOLVED IN NA GLYCOCHOLATE
(Verzár and Kúthy)

No.	pH	Per cent. oleic acid.				Per cent. of fat not diffusible.
		In dialyser.		In outside fluid.		
		Beginning.	End.	Beginning.	End.	
1	6·5	0·081	0·052	0·0	0·042	—
2	6·5	0·250	0·172	0·0	0·078	37·6
3	7·0	0·059	0·033	0·0	0·027	—
4	7·0	0·220	0·142	0·0	0·069	37·3

With the more dilute solutions, the fatty acid was almost completely diffusible, but with the stronger ones, only about two-thirds of it was really dissolved.

It cannot be doubted that the phenomena of hydrotropism may play a very important *rôle* in connection with the absorption of fats, sterols,‡ calcium salts, and other substances which are normally insoluble in water.

The Absorption of Fats

At least 95 per cent. of ingested fat is normally absorbed. The processes of digestion of fat result, first, in very fine emulsification (p. 27), and second, in hydrolysis to glycerol mono- and di-glycerides and fatty acids. For the rapid absorption of fats, three factors are necessary, firstly the fat must be, at least in part, hydrolysed, secondly bile salts must be present ; these two factors provide the trinity of components—monoglyceride, fatty acid, bile salts—which most effectively promotes emulsification ; thirdly the enzyme phosphatase must be present in the gut : if any condition is lacking, absorption is delayed. Analyses of the contents of the gut after a fatty meal show § the emulsion to consist of negatively-charged particles of 0·5μ diameter, or less; some of the fats are present as mono-glycerides and fatty acids, the amount of these as compared with unchanged fat increasing as we descend the gut. Some of the fat may thus be presented for absorption in a state of true solution in consequence of the hydrotropic action of the bile salts. Very little of it can be presented in the form of a soap, since the pH of the contents is about 6·5, but soaps are unstable at reactions more acid than pH 9·0.

---

* NEUBERG. *Bioch. Ztschr.*, 1916, **76**, 105 ; VERZÁR. *Ergebn. d. Physiol.*, 1931, **32**, 421.

† VERZÁR and KÚTHY. *Bioch. Ztschr.*, 1929, **205**, 369 ; **210**, 265, 281 ; 1931, **230**, 451; KÚTHY. *Pflüger's Arch.*, 1930, **225**, 567 ; VERZÁR. *Ergebn. d. Physiol.*, **32**, 391.

‡ GARDNER and GAINSBOROUGH. *Quart. J. Med.*, 1930, **23**, 465.

§ ELKES *et al.* *Proc. R. S.*, 1945, **184A**, 102.

It is probable that some fat is absorbed without being split at all, but the means by which this is effected is unknown. One condition for this seems to be very fine emulsification (particles less than 0·5μ). Even an inert substance like liquid paraffin is absorbed if it is first artificially reduced to an emulsion of a fineness comparable with that to which the fats are reduced in the normal course of digestion.*

If an animal be examined a few hours after a meal rich in fats, the lymphatics of the intestine, the *cisterna lymphatica magna* and the thoracic duct are seen to be distended with a milky fluid—*chyle.* The opacity is due to the presence of minute granules (0·05 to 1 μ) called *chylomicrons.* These consist, not of fatty acids, but of neutral fat.† The fat in such chyle may amount to over 6 per cent., so that in a moderate-sized dog 12 grammes of fat may be carried in the course of an hour from the intestine to the blood by this means.‡ The fact that this is mainly or entirely neutral fat means that, if lipolysis occurred prior to absorption, resynthesis has taken place during absorption. The esters formed are not always glyceryl esters—sometimes cholesteryl esters are found in chyle. Chylomicrons begin to increase in the blood about thirty minutes after a standard meal containing about 30 grammes fat, and reach a peak at about two hours, the peak being higher for obese subjects than for thin ones (Fig. 582). During fat absorption the plasma and serum are therefore milky, and on standing, fat

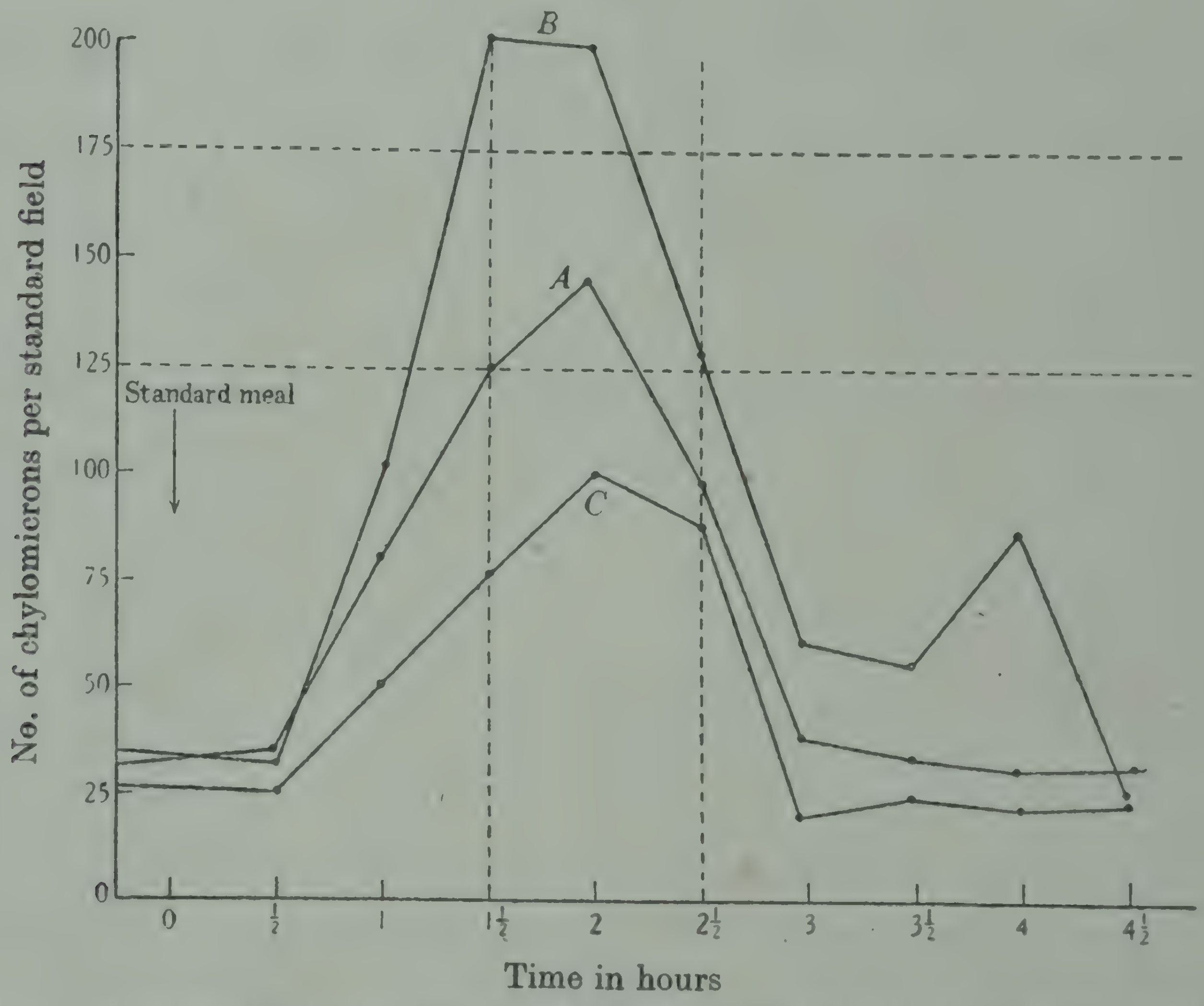


FIG. 582. Counts of chylomicrons (chylomicrographs) of A, normal subject B, fat subject; C, thin subject. Grid marked in dots indicates normal limits. (FRAZER and STEWART. *J. Physiol.*, 1939, **95**, 23P.)

---

* FRAZER *et al.* *J. Physiol.*, 1944, **103**, 306.

† ELKES, FRAZER and STEWART. *J. Physiol.*, 1939, **95**, 68.

‡ GAGE and FISH. *Amer. J. Anat.*, 1924, **34**, 1; FRAZER and STEWART, *J. Physiol.*, 1939, **95**, 21P, 23P.

globules may rise like cream to the surface of the serum. Fat is found in a free state in this finely divided condition in the blood plasma so long as it is being absorbed from the intestine.*

The Blood Lipides. The blood lipides consist of fats, phospholipides and cholesterol. In the post-absorptive state the plasma contains about 600 mg. lipides per 100 c.c. Of this about 100 mg. is usually neutral fat, but this may be absent; phospholipides (calculated as lecithin) amount to about 250 mg./100 c.c.; cholesterol about 150 mg./100 c.c., of which half is free and half is in the form of esters. The red cells contain about 700 mg. lipides per 100 c.c., distributed roughly as follows: neutral fat, 0–150 mg.; phospholipides, 450 mg.; cholesterol (nearly all free), 175 mg.

After a meal containing fat the blood fat and phosphatides show a rise after about two hours, and remain high as long as absorption of fat continues, say for a further two to seven hours. The cholesterol esters of the plasma also increase in fat absorption. Apparently, during absorption more of the lecithin is present in the corpuscles than in the plasma. In from six to ten hours after absorption is complete, free fat disappears, the serum becoming perfectly clear. Thus part at any rate of the fat which is absorbed from the gut is carried thence by the lymphatic channels in the form of neutral fat to the blood stream, by which it is distributed to the various tissues of the body, gradually leaving the blood stream in a manner which at present has not been determined. After a heavy meal of fat, and while the fat content of the blood is high, there appears to be engulfment of fat particles by the endothelial cells of the blood capillaries; the further fate of this fat is uncertain, but it may be passed on into the tissues.

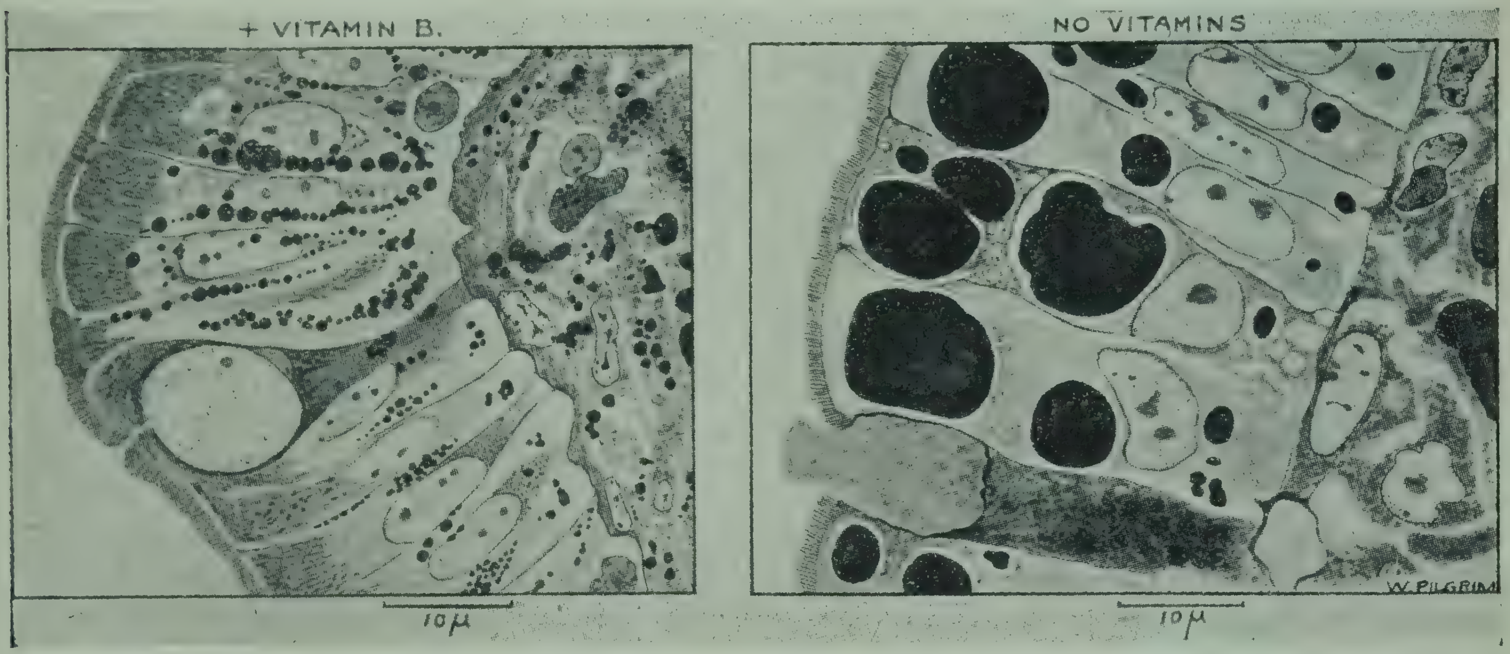


FIG. 583. Epithelium of villus during fat absorption in the rat. On left, showing absorption by streams. On right, absorption by drops. The fat is stained black with osmic acid. (MOTTRAM, CRAMER and DREW, *Brit. J. Exp. Path.*)

Complete starvation causes an increase in all the blood lipides, but the rise in neutral fats is especially great if the subject is obese.

Paths of Fat Absorption. It seems evident that much of the fat which is absorbed takes the path by way of the lymphatics and thus into the blood plasma, though certain facts have presented difficulty in the way of accepting this as a complete view of the process. Ligature of the thoracic duct certainly impedes the absorption of fat, but does not abolish it; the percentage of fat in the blood rapidly falls to a minimum which remains constant, even during starvation. If now fat be administered, although a considerable proportion of it may be absorbed, the percentage of fat in the blood is not raised. If the thoracic duct lymph be collected during the absorption of a given quantity of fat from the intestine, sometimes not more than 60 per cent. of the fat which has disappeared from the gut can be recovered from the lymph. The explanation seems to be that fat is absorbed into both blood and lymph, the fully hydrolysed products largely into blood, the unsplit fat into lymph. If fatty acid or prelipolysed fat is ingested there is lipæmia in portal blood but not in systemic blood, nor much fat in the chyle.* The blood-absorbed fatty acid is taken up from the blood by the liver, the chyle-fat probably by fat depôts.

---

* NISSEN. "Studies in Alimentary Lipæmia in Man." (1933, Copenhagen: Levin and Munksgaard.)

Microscopic examination of a section of a villus during fat absorption shows that absorption occurs through the epithelial cells, in which droplets of fat appear in abundance; they are also demonstrable within the lacteal vessels. We are thus faced with three problems: how does the fat pass into the epithelial cells, what happens to it there, and how does it pass out and into the lacteals? To none of them can an adequate answer be given.

Fatty acids solubilised by the hydrotropic action of bile salts may be provisionally accepted as capable of passing right through the epithelium and into the blood.

It has been credibly suggested that the bile salts are preferentially absorbed, or adsorbed, by the epithelial cells covering the villi, and there exert their solvent action to great advantage. The lecithin of the bile greatly augments the hydrotropic action of the bile salts.† Whereas a 5 per cent. sodium glycocholate solution dissolves 0·12 per cent. of oleic acid, the presence of 0·8 per cent. lecithin raises the solubility to 4·5 per cent.‡ We have seen that the bile salts undergo a circulation, so that the same quantity of bile salts may therefore be used over and over again. That the presence of bile in the intestine is essential for the normal absorption of fat is shown by the fact that on establishment of an external biliary fistula, the utilisation of fat sinks from about 98 per cent. to about 40 per cent., the unabsorbed fatty acid appearing in the fæces.

The absorption of insoluble, unsplit fats, presents greater difficulty. No doubt the cell surface is itself of a lipoidal nature, and fine particles of fat might melt into it: the outer striated hem has also been thought to contain fine channels through which particles might pass.

It has been claimed that absorption of fat through the cells may occur in one of two different ways.§ In the first method, called "*absorption by streams*," a stream of fine fat droplets passes on either side of the nucleus, and the outer part of the cell is free from droplets. Abundant droplets of fat are also seen in the reticular tissue of the core of the villus (Fig. 583).

In the other method, called "*absorption by drops*," the columnar cells are found closely packed with fat granules (Fig. 583) which, small at the beginning of the process of absorption, rapidly enlarge till they occupy the greater part of the cell lying between the nucleus and the striated border. No fat globules are to be seen within the border itself. Mottram, Cramer and Drew found that when vitamins A and B were absent from the food, absorption was by drops. Frazer could not confirm this, but found that addition of choline caused somewhat similar results. It has been suggested that absorption by drops occurs when unhydrolysed fat is absorbed direct.

In the process of fat absorption, the fat is collected in the first place near the Golgi apparatus of the cell, which becomes greatly enlarged during the absorption, and no doubt plays some part in it (Fig. 584); Cramer and Ludford ‖ suggest that it effects the synthesis of neutral fat from fatty acid.

Investigations by special staining methods which distinguish fatty acids from neutral fat or soap ¶ showed that no soap was histologically detectable, but there was plenty of fatty acid, as well as neutral fat, in the epithelial cells, and fatty acid was still found even in the lacteals.

During fat absorption there is some conversion of the fatty acids in the cells to phospholipid compounds by phosphorylation under the action of the intestinal phosphatase. This explains the increase in the phospholipides of the blood during fat absorption; some of the lecithin added to the blood is adsorbed by the red cells. When fat is given together with radioactive "labelled" phosphorus compounds, the phospholipides of the intestinal mucosa in a short time contain the distinctive element. The importance of the enzyme phosphatase is shown by the fact that fat absorption is inhibited by the administration of iodoacetic acid, which poisons this enzyme.

---

* Frazer. *J. Physiol.*, 1943, **102**, 306, 329.

† Moore and Parker. *Proc. Roy. Soc.*, 1901, **68B**, 64.

‡ Fürth and Scholl. *Bioch. Ztschr.*, 1930, **222**, 430.

§ Mottram, Cramer and Drew. *Brit. J. Exp. Path.*, 1922, **3**, 179.

‖ Cramer and Ludford. *Journ. Physiol.*, 1925, **60**, 342.

¶ Weiner. *Ztschr. f. mikr. anat. Forsch.*, 1928, **13**, 197.

How the droplets of fat are passed from the interior of the epithelial cells into the reticular tissue of the villus, and from there onwards into the lumen of the lacteal, remains a mystery. There is definite evidence* that the fat is not transported to any extent, as was once supposed, by leucocytes. It is conceivable that, if the epithelial cells, when sufficiently charged, extruded their droplets into the reticular tissue, they might be transferred thence into the lacteal, together with lymph, by the pumping action of the villi, since we have no positive evidence that the walls of the lacteals are continuous under all conditions. In any case, the contractions of the villi may well play an important part in the process by emptying the lacteals into the submucous plexus of lymphatics, and so on into the valved lymphatics of the mesentery.

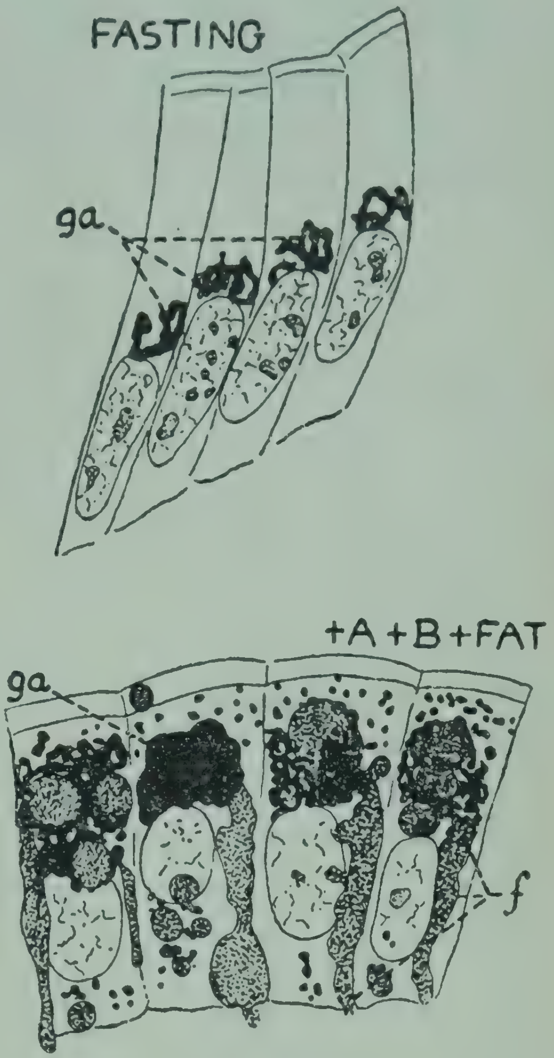


FIG. 584. Showing relation of Golgi apparatus to fat absorption. Above, fasting cells of villus. Below, cells during absorption of fat. *ga* = Golgi apparatus. *f* = fat droplet. (CRAMER and LUDFORD, *J. Physiol.*)

Absorption is interfered with by ligature of the pancreatic duct, so removing lipase from the gut. If the fæces be analysed, it is found that a large proportion of the fat has been split into fatty acids which have remained unabsorbed. This lipolysis has, however, been carried out by the agency of micro-organisms, *i.e.* in the lower segments of the gut where the greater part of the bile has been already reabsorbed. If fat, in a finely divided form, be given to animals deprived of their pancreas, a certain proportion of it is absorbed. Under these conditions some lipolysis may occur in the stomach itself, or absorption of the emulsified fat may have occurred directly.

Substances which are physically almost identical with fats, *e.g.* liquid paraffin, are almost unabsorbed, except when introduced into the intestine in the finest possible emulsion ; after prolonged feeding with liquid paraffin small amounts of it can be found in the liver.† Cholesterol and other sterols, however, are absorbed, most readily when fat is also given, and may be recovered from the liver ; the ease of absorption runs parallel with the solubility of the substance in bile salts.‡

Removal of the whole large intestine does not affect fat absorption.

The Absorption of Carbohydrates

As a result of the action of the various digestive juices, all the digestible carbohydrates of the normal diet of man are reduced to monosaccharides. In the absorption of these by far the greatest part is played by the small intestine. By the time that the food has arrived at the ileocæcal valve, practically the whole of the carbohydrate constituents of the food have been absorbed by way of the vessels of the portal system. The lymph from the thoracic duct contains no more sugar than does the arterial blood taken at the same time, whereas there is an increased percentage of sugar in the portal blood, and even in the general circulation, during the absorption of a big carbohydrate meal.

---

* LEACH. *J. Physiol.*, 1938, **93**, 1.

† CHANNON and COLLINSON. *Bioch. Journ.*, 1929, **23**, 676.

‡ CHANNON and COLLINSON. *Bioch. Journ.*, 1928, **22**, 391.

The problem as to the passage of these substances across the intestinal wall into the blood vessels has already been dealt with. A most striking fact is the relative impermeability of the intestinal wall to the disaccharides as compared with the monosaccharides.

The Absorption of Proteins

The amino-acids produced by protein digestion are absorbed by the blood vessels of the intestinal villi, and pass unchanged into the portal blood. There is no significant increase in the amount of lymph flow or in the amount of protein contained in this lymph during digestion. Moreover, it was shown by Schmidt Mülheim that the absorption of protein was not interfered with as the result of ligature of the thoracic duct.

During the absorption of a protein meal, changes have been described by various observers in the structure of the villus. Several observers have described changes in the epithelium as the result of protein digestion, the cells of the villi becoming swollen and vacuolated.

The first question we have to decide is whether the proteins are always completely hydrolysed before absorption. There is some evidence that serum proteins may be absorbed by the small intestine, though, unless unchanged protein can be shown to have passed into the blood, we cannot exclude the possibility that hydrolysis had occurred owing to the action of intestinal enzymes. An argument in favour of active intervention of the cells of the gut in absorption is perhaps furnished by the fact that if an animal's own serum be introduced into a loop of its intestine the whole of the serum is absorbed. It is not easy to explain this result as due to the digestion of the proteins and their conversion into diffusible products, since the intestinal loops were washed free of any trypsin, and serum has a strong antitryptic action which would prevent its being attacked by traces of trypsin. It has also been claimed by Mills[*] that tissue fibrinogens can be absorbed from the intestine and recovered unchanged from the urine. Foreign proteins, such as egg albumin, or the serum of other species introduced into the blood stream, may cause poisonous effects and albuminuria. If injected in small quantities they excite, as a reaction on the part of the organism, the production in the blood serum of a *precipitin*, and later the animal becomes anaphylactically sensitised (Chapter IV.), and the presence of the precipitin or of anaphylactic sensibility may therefore be looked upon as a test by which we may decide whether these proteins have passed through the intestinal wall unchanged. In most cases it is found that, however abundant the amount of protein administered, none of it appears in the urine, nor is any precipitin formation or anaphylactic reaction aroused. Ascoli, however, has observed such results occasionally to follow the administration of large doses of egg white, and it has been shown[†] in the sensitised human subject that "allergic" reactions may then ensue when the specific native protein is eaten, or even given per rectum. In experiments on guinea-pigs this was not the case unless there was simultaneous damage to the intestinal wall, or considerable pressure applied within the gut.[‡] There is a difference, according to age, in the behaviour of animals to the introduction of soluble protein into their alimentary canal. It seems that during the first few days of life the cellular lining of the alimentary canal is slightly permeable to foreign proteins, whereas later on little or no protein is absorbed unchanged from the gut.

The absorption of unchanged proteins can therefore play but a small part in the assimilation of protein as a whole.

When peptones and proteoses are introduced into the intestine they are rapidly broken down to amino-acids by enzymes of the intestinal juice ; this explains why, in the earlier experiments when peptone was absorbed no peptone could be found in the portal blood—a fact which led to the incorrect deduction that the peptone had, in the course of absorption, been synthesised to form native proteins of the blood plasma.

We know from the researches of Fischer and others that, although the different proteins in our food present a marvellous qualitative similitude, in that all of them yield on hydrolysis the same kinds of amino-acids, there are great differences in the relative amounts and stero-chemical arrangement of these amino-acids in different proteins. It is difficult to imagine how, for instance, muscle protein could be formed from wheat protein, a process continually occurring in the growing animal, unless we assumed that the protein molecule is first partly taken to pieces, and that its constituent molecules are then selected by the growing cells of the body and built up in the order and proportions which are characteristic of muscle protein.

---

[*] MILLS. *Amer. J. Physiol.*, 1923, **63**, 484.

[†] SÜSSMANN, DAVIDSON and WALZER. *Arch. Int. Med.*, 1928, **42**, 409 ; BRUNNER and WALZER. *Arch. Int. Med.*, 1928, **42**, 172.

[‡] HETTWER and KRIZ. *Amer. J. Physiol.*, 1925, **73**, 539.

Very little energy is lost in such transformation, *i.e.* the heat of combustion of the products of digestion would differ from that of the original protein by less than 5 per cent.

Since all the evidence shows that ingested protein undergoes complete hydrolysis before entering the blood stream, it should be possible to feed an animal on a diet in which the protein had been replaced by the corresponding mixture of amino-acids. The experimental decision of this question was first attempted by O. Loewi,* who kept a dog in a state of nitrogenous equilibrium on a diet containing fat, starch, and a pancreatic digest of protein, which contained no substances giving the biuret test. These results have been confirmed,† and it has also been found that the intravenous injection of mixed amino-acids will keep a man in nitrogenous equilibrium if equal in N-content to the requisite amount of protein ‡; the administration (plus glucose), of 2·5 to 10 per cent. enzymic protein hydrolysates, or of acid hydrolysates to which tryptophan and methionine have been added, is now a recognised clinical procedure in dealing with almost moribund patients too weak to be fed.§ For starved and declining prisoners in concentration camps in 1945, however, oral administration proved as good as injection.

The Fate of the Amino-Acids after Absorption. During starvation the protein requirements of the active tissues are met at the expense of the less active tissues. The proteins characteristic of any tissue can be taken down, removed to another part of the body, and built up into the proteins characteristic of some other active tissue. It is difficult to conceive that such a transference and transformation could occur in any other way than by a more or less thorough disintegration of the protein molecule at one place and its synthesis at the other, and we know that every tissue contains intracellular enzymes (cathepsins) which are capable of effecting the disintegration of proteins, and are responsible for the autolysis of tissues after death. There is thus no *a priori* reason to believe that the blood carries the proteins from the alimentary canal to the tissues in any other form than that of amino-acids. Strongly in support of this conclusion is the evidence ‖ that after a protein meal there is an increase in the non-protein nitrogen (*i.e.* the nitrogen not present as coagulable protein) of the blood, and particularly in the portal vein blood. The final proof was rendered possible by Van Slyke's method ¶ for the detection of small quantities of amino-acids in the blood and tissues, and ultimately by the actual isolation of amino-acids from the dialysate prepared by circulating blood from a living animal through a collodion dialyser,** and back to the animal again. Several different amino-acids have been isolated from the blood,†† though it is not known to what extent or under what circumstances the individual amino-acids are altered in amount. In sum, however, their amount does vary. The blood always contains a certain amount of amino-acids, even during fasting, but after a protein meal there is a considerable increase in the amount.‡‡ Thus the blood of fasting animals contains from 3 to 6 milligrammes of amino-acid nitrogen per 100 c.c. Blood taken after food contains from 8 to 10 milligrammes of amino-acid nitrogen per 100 c.c.

---

* Loewi. *Arch. f. exp. Path.*, 1902, **48**, 303.

† Henderson and Dean. *Amer. J. Physiol.*, 1903, **9**, 386.

‡ Martin and Thompson. *Medicine*, 1943, **22**, 73.

§ Beattie *et al.* *B.M.J.*, 1944, *i.*, 209.

‖ Cathcart and Leathes. *J. Physiol.*, 1905, **33**, 462; Folin and Denis, *J. Biol. Chem.*, 912, **11**, 87; **12**, 141.

¶ Van Slyke. *J. Biol. Chem.*, 1911, **9**, 185 (original form); *ibid.*, 1929, **83**, 425 (improved form). A colorimetric method has been given by Folin, *J. Biol. Chem.*, 1922, **51**, 377.

** Abel, Rountree and Turner. *J. Pharmacol.*, 1914, **5**, 275.

†† Abderhalden. *Z. f. Phys. Chem.*, 1913, **88**, 478.

‡‡ Van Slyke and Meyer. *J. Biol. Chem.*, 1913, **16**, 197, 213, 231.

In Fig. 585 are shown the results of experiments on rats absorbing a diet of meat. The upper curve (I) shows the per cent. of absorption from the intestine; the lower curves show (II) the amino-nitrogen of the blood, and (III) the oxygen consumption of the animal. During the first three hours absorption is rapid, the amino-N of the blood rises steeply, and with it the oxygen usage—a similar increase of oxygen usage is obtained if amino-acids are injected direct into the blood (specific dynamic action).

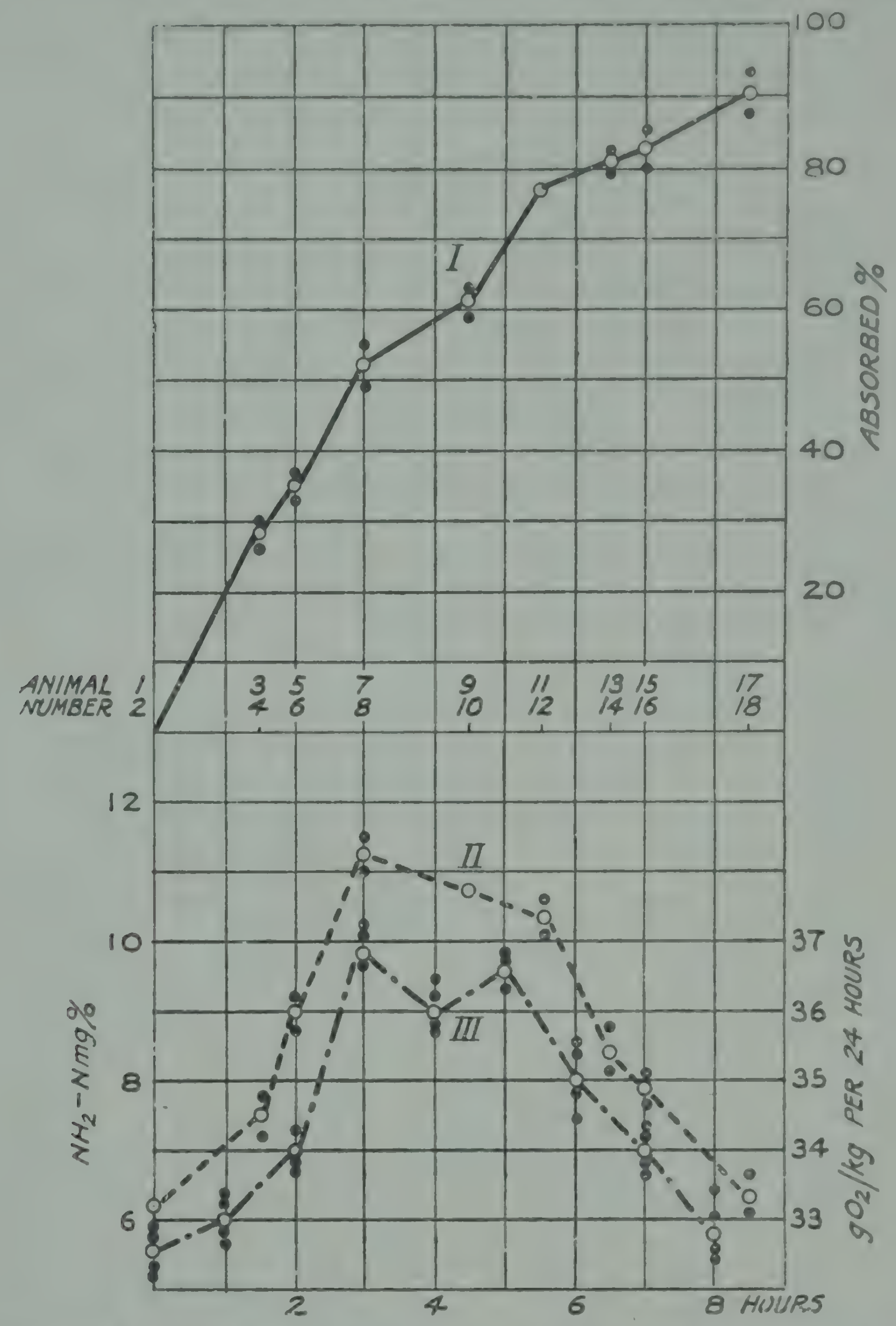


Fig. 585. Effect of feeding rats with meat. I. Rate of absorption from alimentary canal. II. Amino-N of blood. III. Oxygen usage of animals. (Kúthy.) (From Verzár, *Ergebn. d. Physiol.*, **32**, 1931.)

The question of the fate of amino-acids thus absorbed from the intestine into the blood is decided by an estimation of the amino-acid content of the blood and of different tissues at various times after the injection of amino-acids into the blood. After the injection of amino-acid, only a certain proportion is excreted in the urine, and the rest of the amino-acid rapidly disappears from the blood and is taken up by the tissues without undergoing any immediate chemical change: in certain tissues, such as the muscles, a definite saturation point exists which sets the limit to the amount of amino-acids that can be absorbed, but the capacity of the internal organs, and especially of the liver, for the absorption of amino-acids is much greater.

It is worthy of note, however, that the abstraction of amino-acids by the tissues from the blood is never complete, *i.e.* the amino-acids of the blood must be in a state of equilibrium with those of the tissues, although the concentration in the latter may be much greater than in the former. If several hours be allowed to elapse after the injection of amino-acids before the analysis of the tissue is undertaken, it is found that the amino-acid nitrogen content of the liver may have returned to normal, although the concentration in the muscles has suffered no appreciable fall. We know that this difference is due to the fact that the liver is the main site of amino-acid breakdown.

We therefore conclude that the amino-acids produced by protein digestion are absorbed without further change into the blood stream. They then circulate throughout the body, as a result of which an equilibrium concentration is established for each amino-acid between the blood and the several tissues, a certain proportion of them finally being built up in each tissue into the proteins characteristic of that tissue in order to replace the waste caused by wear and tear. The rest, probably the major part of the amino-acid, is taken up by the liver, where most of it undergoes deamination, the nitrogen moiety being rapidly converted into urea which is excreted by the kidneys, while the non-nitrogenous moiety is carried to the working tissues, which it serves as a ready and immediate source of energy.

The distribution of the amino-acids in fasting, starving and during absorption has been studied in cats by Bolton and Payling Wright.* They estimated the amino-nitrogen in blood and in plasma coming from the intestine by the superior mesenteric vein, in that leaving the liver by the hepatic vein, in blood going to the tissues by the carotid artery, in blood leaving the tissues by the inferior vena cava, and in the lymph.

Some of their average results are given below :—

Condition.	Average amino-N (mg/100 cc.) in blood plasma of				Lymph.
	Carotid.	Sup. Mes Vein	Hepatic vein	Inf. ven. cava	
Fasting . . .	5·0	5·3	4·8	5·3	5·9
Starving . . .	4·9	4·9	4·4	5·4	6·1
Fed peptone 20 g. 5 hrs. before . . .	15·2	19·5	15·9	14·9	17·1

These experiments bear out the current view that absorption of the amino-acids into both the capillaries and the lymphatics of the villi is in accordance with the laws of diffusion. The blood capillaries being more superficial, and the lymph being diluted by the tissue fluid derived from the afferent blood stream, the amino-acids reach a higher level in the portal blood than in the lymph during absorption. Always there is a lower level of amino-acids in the hepatic vein blood, showing that the liver is constantly removing amino-acids from the blood. Also, in fasting, the blood from the tissues and intestine has the same content of amino-acid, while in starvation the venous return from the tissues has a raised content as compared with the arterial blood, which shows that the muscles and other tissues are undergoing a process of breakdown, yielding amino-acids to the blood and lymph.

The fact that not only the blood but also the tissues contain amino-acids, even after complete starvation for some days, shows that these substances are intermediate steps not only in the synthesis but in the breaking down of body proteins. Free amino-acids are thus the protein currency of the body, just as glucose is the carbohydrate currency. The fact that high protein feeding does not appreciably increase the amino-acid content of the tissues shows that any storage of nitrogen in the organism must take place, not in the form of amino-acids, but as tissue protein.

The Plasma Proteins. The plasma proteins, serum albumin and serum globulin, probably play an important *rôle* in protein metabolism. The plasma proteins are almost certainly formed, in the liver, from the amino-acids derived from digestion of protein ; they then act as a reserve of readily available protein in case of need. In prolonged protein starvation, the plasma proteins are lowered, and when, in man, the level drops below 5·5 per cent., hunger œdema may occur, because of the lowered colloidal osmotic pressure of the plasma.

---

* Bolton and Payling Wright. *J. Physiol.*, 1937, **89**, 269.

The plasma proteins can be experimentally lowered by the process called *plasmapheresis*. In this, blood is drawn daily, and the corpuscles washed with isotonic salt solution and reinjected. By suitable, steadily diminishing, daily plasmaphereses the serum proteins can be lowered, say, from 6 per cent. to a certain low level, say, 4 per cent. The plasma protein that must be removed per week in order to maintain this level reaches a constant figure, called the basal output. If anything is done to upset the balance, more or less plasma protein must be removed in order to keep the constant low level. Feeding with different proteins, for instance, will alter it. Thus, 100 grammes ox serum protein will form 38 grammes serum protein in a dog, 100 grammes meat forms about 20 grammes, and 100 grammes gelatin only 5 grammes.

Plasma protein can be formed also from reserve stores in the body, and fasting animals can replace 40 to 60 per cent. of the circulating plasma protein within twenty-four hours, and even after an extensive hæmorrhage up to 25 per cent. of the original plasma protein is reinstated in about twelve hours. The plasma proteins so added are apparently not drawn from a preformed stock, but are constructed within a few hours.* Conversely, plasma protein can be utilised for general purposes. Thus a protein-starved dog can be kept in nitrogenous equilibrium by frequent injections of dog plasma. This accession of protein appears to benefit the tissues directly, for it is not followed by an increase of nitrogen output as dietary protein intake would be, nor does it raise the blood amino-acids, although the injected plasma protein has all disappeared in less than twenty-four hours. Starved people in prison camps benefited by intravenous serum injections.

It seems probable, therefore, that the tissue proteins are supported by amino-acids and/or plasma proteins. Part of the tissue protein forms a reserve and is readily and quickly depleted. Depletion occurs in protein starvation, illnesses due to infections, burns, fractures, etc. Supplies of amino-acid from without, and demand for protein within, are in a condition of dynamic equilibrium with protein reserves. If the need for protein for cells or plasma exceeds the supply from amino-acids, the reserve stores, including plasma proteins, will be depleted, and if less it will be replenished. It seems probable that tissue proteins can be built from plasma proteins without loss of nitrogen.

THE ACTUAL COURSE OF DIGESTION. London † described the course of digestion of meals in dogs which had been provided with fistulæ at various points along the alimentary canal.

We may take as an example the digestion of 200 grammes of bread. This evoked the secretion of 20 grammes of saliva, about 10 grammes of mucus from the coats of the stomach, and about 315 grammes of gastric juice. In the animal with a pyloric fistula, the first admixture of the food with the outflow from the fistula occurred at eight to twelve minutes after the completion of the meal, and after this time the pylorus continued to open at regular intervals, discharging each time a small amount of fluid composed of particles of undigested bread mixed with gastric juice. The following Table shows the percentage amount of food taken which had left the stomach at the end of each hour after the meal :

First hour	32·6 per cent.
Second hour	17·9 ,,
Third hour	29·5 ,,
Fourth hour	1·87 ,,
Fifth hour	6·66 ,,
Sixth hour	4·21 ,,

The outflow from the duodenum, being mixed with the pancreatic juice and the bile, is yellow in colour and increased in amount. With a meal of 200 grammes there are secreted on the average 130 grammes of bile and 140 grammes of pancreatic juice.

---

* EBERT *et al.* *Amer. J. Physiol.*, 1942, **136**, 299.

† LONDON. *Z. f. Physiol. Chem.*, 1907, **53**, 246 ; 1908, **56**, 378 ; 1909, **60**, 191, 267.

On deducting the amount of juices which have been added to the food, it is found that even in this short length of intestine absorption of about one-sixth of the ingested food has taken place, about one-fourth of the carbohydrates having been absorbed and about one-eighth of the proteins.

In a dog with a fistula about the middle of its small intestine, the material obtained was quite different in character from that flowing from the duodenal fistula. The pasty character had disappeared, the material forming a frothy, orange-yellow, even jelly-like mass with practically no trace of undigested bread.

From a fistula in the ileum the outflow was much scantier than that obtained from the jejunal fistula, consisting of a thick jelly-like, orange-coloured mass. Both proteins and carbohydrates were entirely digested.

From a fistula in the cæcum there was no outflow until four or five hours after the meal. The material from the gut was then extruded in fæcal-like masses at intervals of one half to one hour. This regular outflow lasted for about six hours. The reaction of the contents was strongly alkaline, with no food particles, and the material contained merely débris of cells, with small traces of sugar, dextrin and unaltered starch. The absorption of the foodstuffs is thus practically complete by the time that the food has reached the lower end of the small intestine.

The following Table gives the total amounts obtained in a series of experiments from the different fistulæ after administration of 200 grammes of bread, and also the percentage amount of foodstuffs which had been absorbed at each level :

	Total amounts obtained from 200 grms. of bread.	Absorbed per cent.
Pyloric fistula	691 grammes	0
Duodenal fistula	691 ,,	17·45
Jejunal fistula	585 ,,	37·77
Ileal fistula	412 ,,	67·65
Cæcal fistula	80 ,,	94·34

For Reference

Bloor (1939). Fat Transport in the Animal Body. *Physiol. Rev.*, **19**, 557.

Fischer (1947). The Participation of Blood Proteins in the Metabolism of Tissue Cells. *Biol. Rev.*, 1947, **22**, 178.

Goldschmitt (1921). On the Mechanism of Absorption from the Intestine. *Physiol. Rev.*, **1**, 421.

Karel (1948). Gastric Absorption. *Physiol. Rev.*, **28**, 433.

London (1913). Physiologische u. pathologische Chymologie. (Akad. Verlagsgesselschaft, Leipzig.)

Madden and Whipple (1940). Plasma Proteins. *Physiol. Rev.*, **20**, 194.

Magee (1930). The Rôle of the Small Intestine in Nutrition. *Physiol. Rev.*, **10**, 473.

Verzár (1936). Absorption from the Small Intestine. (London : Longmans Green.)

CHAPTER XLVI

THE HISTORY OF THE FOODSTUFFS

The absorbed products of the foodstuffs have varied histories in the body; for a while some of them, after modification, may linger on in new guises as component parts of the living framework, but sooner or later all of them suffer ultimate oxidative breakdown to relatively simple compounds, which are got rid of by excretion. The physiological history of these chemical transformations is called *metabolism*, or *intermediary metabolism*.

For purposes of description the subdivision of the subject under the headings of protein, carbohydrate or fat metabolism is retained. But it should be realised that such subdivision is largely artificial, and that actually these various metabolic highways have crossroads in common, and so are interrelated. Moreover, the distinction between the materials of the living machinery and the non-living fuel which it uses are less sharp that was once supposed.

The Use of Isotopes in the Study of Metabolism

An important aid in following the fate of a substance in the body is provided by the use of elements in isotopic form. There are two types of isotopes which may be used. The first of these are the naturally occurring ones.

When an element occurs in both normal and isotopic forms, these differ only in atomic weight, and physical properties which depend on it, but are otherwise indistinguishable. The proportion of an isotope in a given element is constant, whether the element be free or in combination. Thus nitrogen, whether in air, in an amino-acid, ammonia, urea or protein, always contains 1 part of N^{15} (*i.e.* nitrogen of atomic weight 15) to 269 parts of the normal nitrogen N^{14}. Similarly, heavy hydrogen, or deuterium (D), of atomic weight 2, occurs with normal hydrogen in the proportion of 1/6,000, and heavy oxygen, O^{18} with O^{16} in the ratio 1/530, and heavy carbon C^{13} with C^{12} in the ratio of 1 : 90. The reason for this is that none of the agencies of nature can discriminate between the mixtures of elements and isotopes. Depending on differences in physical properties, however, an artificial separation can be effected, and so the element or its compounds can be prepared which will contain the isotope in a higher proportion than that found naturally. The amount of an isotope in such an element or compound is capable of determination with accuracy by physical methods, such as determination of density, mass spectrographic analysis, etc. Thus the concentration of N^{15} in nitrogen can be determined to within 0·003 per cent.

As an example of the use of natural isotopes we may mention the employment of N^{15} in the study of the fate of the amino-group of amino-acids. When amino-acids, the nitrogen of which has been "labelled" by the addition of the N^{15} compound, are given with the diet, most of the nitrogen is recoverable in combination in the tissue proteins, and some in the excreta.

The other type of isotope is the artificially prepared radioactive form of the element; thus carbon, phosphorus, calcium, sulphur, sodium, iron, iodine, etc., can be rendered radioactive, and, since the active element can readily be identified and estimated by this property,* the element, whether free or in a compound, can be traced through the body. These forms of elements are also biologically indistinguishable from the normal forms.

As an instance of the use of a radio-active element, we may quote experiments with labelled phosphorus, prepared from sulphur by neutron bombardment. When this was injected subcutaneously into rabbits, as sodium phosphate, it was found that within twenty-seven days 45 per cent. of it was excreted into the urine, and 11 per cent. in the fæces, and the element was shown to be present even in the calcium phosphate of the bones for a few weeks.

---

* Ambrosen *et al.* *Acta Physiol. Scand.*, 1945, **10**, 195.

For Reference

BUCHANAN and HASTINGS (1946). The Use of Isotopically Marked Carbon in the Study of Intermediary Metabolism. *Physiol. Rev.*, **26**, 120.

CHAIKOFF (1942). The Application of Labelling Agents to the Study of Phospholipid Metabolism. *Physiol. Rev.*, **22**, 291.

SCHOENHEIMER and RITTENBERG (1940). The Study of Intermediary Metabolism of Animals with the Aid of Isotopes. *Physiol. Rev.*, **20**, 218.

SCHOENHEIMER (1942). The Dynamic State of Body Constituents. (Harvard : Univ. Press.)

Study of the Metabolism of Tissue Slices

In the study of the metabolic changes which occur in tissues, much may be learnt by the study of tissue slices. Slices of fresh tissues weighing 20–100 mg., and of about 0·3 mm. thickness, cut by hand with a razor, continue to live for some time if kept supplied with oxygen. This is done by placing the tissue slices in Ringer's solution kept supplied with oxygen by shaking. It is easy to determine the rates of oxygen utilisation and CO_2 production of such tissue slices by the use of the Warburg manometer, one pattern of which is shown in Fig. 586.

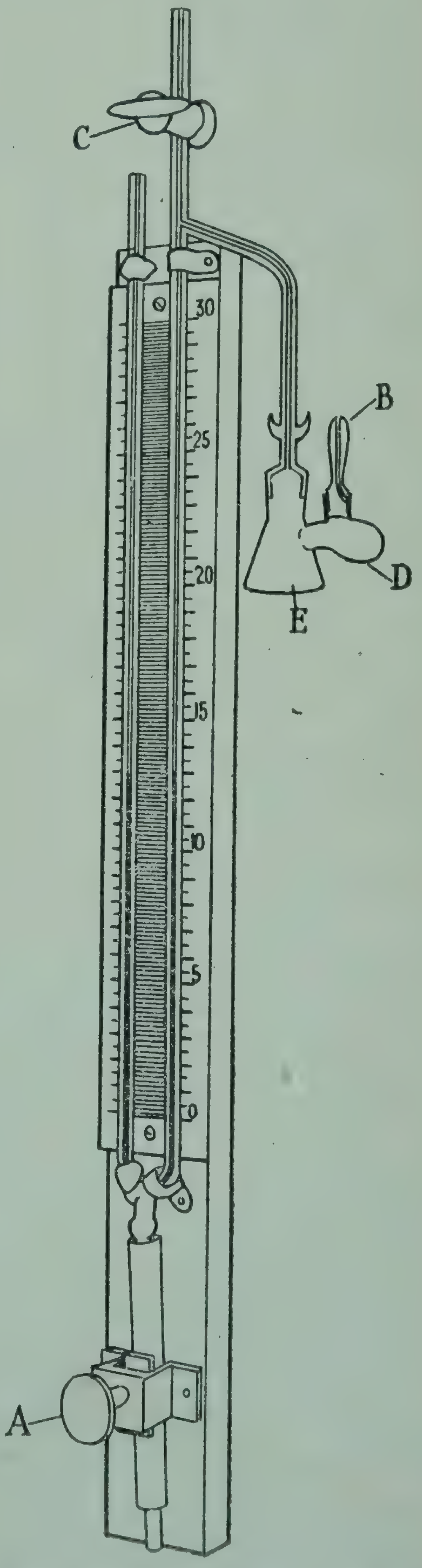


FIG. 586. Diagram of the Warburg type of manometer.

This consists of a long narrow U-tube graduated in millimetres ; one arm of the U is open to the atmosphere, the other leads to a tap C and a flask E of special construction as shown in the figure. The manometer is filled with water containing some bile salts, and opens at the bottom of the U into a rubber reservoir provided with a screw arrangement A for adjusting the amount of liquid in the U-tube. The manometer is to measure changes in the amount of gas contained in the flask and other of the closed parts of the apparatus ; these changes are measured, not by observing alterations in volume at constant pressure, but by noting alterations in pressure at constant volume. For this purpose the liquid in the U-tube starts off with a certain reading on the two sides (*e.g.* they may be at the same level). If gas is evolved or absorbed in the flask this will tend to push the liquid down on the proximal limb and up in the distal limb, or *vice versa* ; the reservoir screw is so adjusted by adding fluid to or abstracting it from the manometer, as to bring the reading in the proximal limb to the level at which it began. If gas has been absorbed, for instance, fluid would have to be withdrawn from the manometer in order to expand the gas in the flask to its initial volume ; if gas has been evolved, fluid would have to be added to the manometer to compress the gas in the enclosed space. By the application of suitable factors and corrections the amount of gas evolved or absorbed can then be calculated.

In use, the flask is filled with a gas of known composition, *e.g.* oxygen, or oxygen and CO_2, by passing oxygen in at the tap B and out by the tap

C. These taps are then closed, so confining the gas within a closed space. The apparatus is then placed in a thermostat at suitable temperature and shaken in a horizontal direction to and fro by a suitable motor. Readings of the pressure changes are then taken at suitable intervals of time.

The tissue slices are placed in the lower part of the flask, and are suspended in Ringer's solution which, as the result of the shaking, soon becomes saturated with oxygen ; subsequent oxygen absorptions then represent oxygen utilisation by the tissues. The CO_2 evolved by the tissue is absorbed by a little KOH solution placed in the inner receptacle F. The side tube D is made to contain any reagent which it may be desired to add to the contents of the flask at any subsequent stage, without having to open the apparatus to the air.

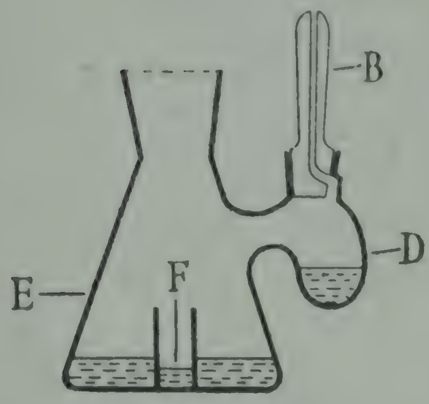


FIG. 586A.

By micro methods of analysis of the Ringer's solution and of the tissue slices many chemical processes can be followed ; by slight modification of the technique the evolution of CO_2 as well as the absorption of oxygen can be determined, and in general a valuable picture may be obtained of the kind of metabolic processes which such surviving slices of tissue are capable of carrying out. It is true, of course, that any such findings cannot be applied without some reservation to explain all the changes which may take place in the intact organ in the body, but the possibilities which the tissue slice technique provides have certainly been found to furnish valuable information.

The results are frequently expressed in terms of so-called Q values which are given in cubic millimetres of gas, *e.g.* Q_{O_2} means the cubic millimetres of oxygen utilised per milligramme dry weight of tissue per hour. Similarly Q_{CO_2} would be the volume of CO_2 similarly evolved. Absorptions, *e.g.* oxygen usages, are given a negative sign and evolutions of gas a positive sign.

1. PROTEIN METABOLISM

The protein of food has a twofold destiny. Of the absorbed amino-acids, a part, usually the greater, is broken down and yields energy, while the smaller part ultimately forms an indispensable constituent of the living substance. This part cannot be replaced by other foodstuffs, and we will consider its history first.

A. Protein Synthesis

The formation of the proteins of the living protoplasm from the proteins of the food is no question of mere translation of a substance from one place to another. For the ingested proteins are broken down into their constituent amino-acids before being absorbed into the blood stream. Furthermore, the tissue proteins are different from those which are ingested, varying from tissue to tissue and from one species to another.

It has been shown that amino-acids are abstracted from the blood by the tissues. In general terms, the synthesis of protein must be effected by the action of enzymes, such as tissue cathepsins, which bring about condensation between different amino-acids. But we have no idea how, under the influence of genes for instance, the constituent amino-acids are assembled always in the correct order and proportions, to produce proteins characteristic of the tissue and of the animal species concerned. The essential amino-acids must be supplied in the diet ; the non-essential ones can, if necessary, be synthesised in the body, so that finally the jig-saw molecule may be duly built up.

A method of studying tissue protein synthesis is that of "repletion" in depleted animals. Rats are given for two to three months a diet low in protein, but otherwise adequate. They lose weight and show hypoproteinæmia. If the protein, or all its essential amino-acids, is now restored there is rapid resynthesis of tissue protein and weight is regained. If only one of the essential amino-acids, *e.g.* lysine is omitted, the resynthesis does not take place. For adequate protein synthesis the total calorie value of the diet must be above a certain minimum, but above this level fat is laid down. Of the vitamins necessary for protein synthesis, riboflavin seems the most essential. This method also indicates that for tissue protein synthesis to occur all the essential amino-acids must be available simultaneously, since no stores of these are carried in the body.

By administration of marked amino-acids, in which the N is N^{15} and one of the H atoms is replaced by deuterium, new light has been shed on the fate of amino-acids in the body.* Much of the added acid is found to be present in protein, and there is reason to suppose that the new amino-acid has been introduced into the protein molecule in exchange for an identical molecule which leaves it. Continuous deamination and re-amination of the amino-acids in protein is also indicated by the use of this method. The proteins of the tissues thus appear to be highly labile and dynamic and not mere static compounds, as once supposed.

The Synthesis of Amino-Acids. The body can not only effect deamination of an amino-acid, but, by reversal of the enzyme action, can synthesise an amino-acid from its corresponding ketonic acid and ammonia. Knoop † found that the next higher homologue of phenyl alanine, namely phenyl-α-amino-butyric acid, when administered to an animal, was excreted in large quantities in the urine as an acetyl derivative, which was easily isolated in a state of purity. The administration of phenyl-α-keto-butyric acid led to the excretion of the same acetyl derivative in the urine. Since keto-acids are ordinary products of the breakdown of amino-acids, it is evident that the animal body can add ammonia and form amino-acids from them.

If the phenyl derivative of the corresponding hydroxy-acid, α-hydroxy-butyric acid, was administered, this was also excreted as the acetyl amino-compound. This and other evidence suggests that amino-, keto- and hydroxy-acids are interconvertible in the body, the connecting link being the keto-acid, *e.g.*

$$\underset{\textbf{Alanine}}{CH_3.CHNH_3COOH} \rightleftarrows \underset{\textbf{Pyruvic Acid}}{CH_3CO.COOH} \rightleftarrows \underset{\textbf{Lactic Acid}}{CH_3CHOH.COOH}$$

These changes can be effected by the perfused liver as readily as by the whole body.‡

The keto- or hydroxy-acids need not be derived from proteins at all but, like pyruvic acid, may result from carbohydrate metabolism. Thus, part of the nitrogen set free by protein disintegration might be recombined with the formation of amino-acids, without appearing in any way in the nitrogen balance-sheet of the body.

Glycine and alanine can be easily manufactured in the body, and the organism is capable of synthesising amino-acids from ammonia and the corresponding keto-acid. For these reactions the ammonia is derived from glutamine. Further, by transamination, one amino-acid can be converted into another.

Transamination.§ This is the name given to a reaction in which a $CHNH_2$ group of one compound can be exchanged for a keto-group of another. One of the bodies must be a *dicarboxylic* α-amino-, or α-ketonic acid (*e.g.* glutamic, aspartic, oxaloacetic or α-ketoglutaric acid) and the other an α-amino or α-ketonic acid (*e.g.* alanine, pyruvic acid). The following is an example :—

$$\underset{\textbf{Glutamic Acid}}{COOH.CH_2.CH_2.CHNH_2.COOH} + \underset{\textbf{Pyruvic Acid}}{CH_3.CO.COOH} \rightleftarrows$$

$$\underset{\alpha\textbf{-Ketoglutaric Acid}}{COOH.CH_2.CH_2.CO.COOH} + \underset{\textbf{Alanine}}{CH_3.CHNH_2.COOH}$$

---

* Schoenheimer *et al.* *J. Biol. Chem.*, 1939, **130**, 703.

† Knoop. *Z. f. Physiol. Chem.*, 1910, **67**, 489 ; 1914, **89**, 151.

‡ Embden and Schmitz. *Biol. Ztschr.*, 1910, **29**, 423 ; 1912, **38**, 393.

§ Braunstein. *Nature*, 1939, **143**, 609.

The reaction is catalysed by enzymes called *aminophorases* or transaminases (in which the vitamin pyridoxol is the prosthetic group), and is an important means by which one amino-acid can be converted into another. It has also been found that, by the subsequent action of dehydrogenases, with co-enzyme I α-ketoglutaric acid and NH_3 can be reversibly converted into glutamic acid while at the same time a donator of hydrogen is oxidised. Glutamic acid may thus exert a catalytic effect in the oxidation of amino-acids to keto-acids and NH_3, as mentioned above.

For Reference

Alcock (1936). The Synthesis of Proteins in Vivo. *Physiol. Rev.*, **16,** 1.

Braunstein (1947). Transamination and the Integrative Functions of the Dicarboxylic Acids in Nitrogen Metabolism. *Adv. in Protein Chemistry*, Vol. III, p. 2. (Acad. Press Inc., N.Y.)

Cannon (1948). The Problem of Tissue Protein Synthesis. *Feder. Proc.*, **7,** 391.

Herbst (1944). The Transamination Reaction. *Adv. in Enzymol.*, **4,** 75.

Wasteneys and Borsook (1930). The Enzymatic Synthesis of Protein. *Physiol. Rev.* **10,** 110.

B. Protein Catabolism

In the body, proteins ultimately become broken down into simple molecules and are excreted. Of these end-products the most characteristic are compounds of nitrogen, and of sulphur, which are almost entirely eliminated in the urine.

Animals are divisible into three groups, according to the form in which they excrete most of their nitrogenous end-products. The first group excretes it mostly as ammonia, and is called *ammonotelic*, the second, called *uricotelic*, excretes chiefly uric acid, and the third, or *ureotelic*, excretes urea. Each group may excrete smaller amounts of nitrogen in the other two forms. The explanation of the distinction, according to Baldwin, is that ammonia is highly toxic, but very diffusible, and so can only be used as an excretory form in aquatic invertebrates. When the supply of water is liable to fluctuation the nitrogen must be handled in the less toxic form of urea, as in mammals, or uric acid, as in birds and terrestrial invertebrates.

One of the facts which is revealed by urine analysis is that the distribution of nitrogen in the form of the various substances differs according to whether the subject is on a high or low protein diet. Not only is there a larger total nitrogen output when the subject is on a high protein diet, but we also find evidence of a qualitative difference as in the Table given here.* In the first place, in a normal protein diet the urea accounts for 87 per cent. of the total nitrogen of the urine. On an excessive protein diet this percentage may rise to

Distribution of Nitrogen in Urine on Various Diets

	July 13 Ordinary diet	July 20 Low protein diet
Vol. of urine	1170 c.c.	385 c.c.
Total nitrogen	16·8 grm.	3·60 grm.
Urea	14·70 grm. = 87·5 %	2·20 grm. = 61·7 %
Ammonia	0·49 grm. = 3·0 %	0·42 grm. = 11·3 %
Uric acid	0·18 grm. = 1·1 %	0·09 grm. = 2·5 %
Creatinine	0·58 grm. = 3·6 %	0·60 grm. = 17·2 %
Undetermined	0·85 grm. = 4·9 %	0·27 grm. = 7·3 %
Total SO_3	3·64 grm.	0·76 grm.
Inorganic SO_3	3·27 grm. = 90·0 %	0·46 grm. = 60·5 %
Ethereal SO_3	0·19 grm. = 5·2 %	0·10 grm. = 13·2 %
Neutral S	0·18 grm. = 4·8 %	0·20 grm. = 26·3 %

---

* Folin. *Amer J. Physiol.*, 1905, **13,** 45, 66, 117.

90 or 95. On a low protein diet, not only is the total amount of nitrogen in the urine much less, but the percentage of nitrogen appearing as urea is reduced to 60. On the other hand, ammonia suffers little absolute diminution, so that its percentage of total N is considerably increased. Also, practically identical *amounts* of creatinine are obtained under the two conditions.

The interpretation which was placed upon these facts was that on high protein diets much of the nitrogenous bodies of the urine were derived almost directly from the food, while on low protein intake a relatively greater part of the substances came from the breakdown of tissue protein. Thus there were two sources of nitrogen of the urine, *exogenous, i.e.* that from the food, and *endogenous, i.e.* that derived from the tissues, but we must regard the implied classification of the nitrogenous derivatives as of the nature of a convenient fiction, in view of the facts which have since been brought to light, and which show that there is a continual interchange of amino-acids, or of atomic groups, between the tissue proteins and the fluids of the body.

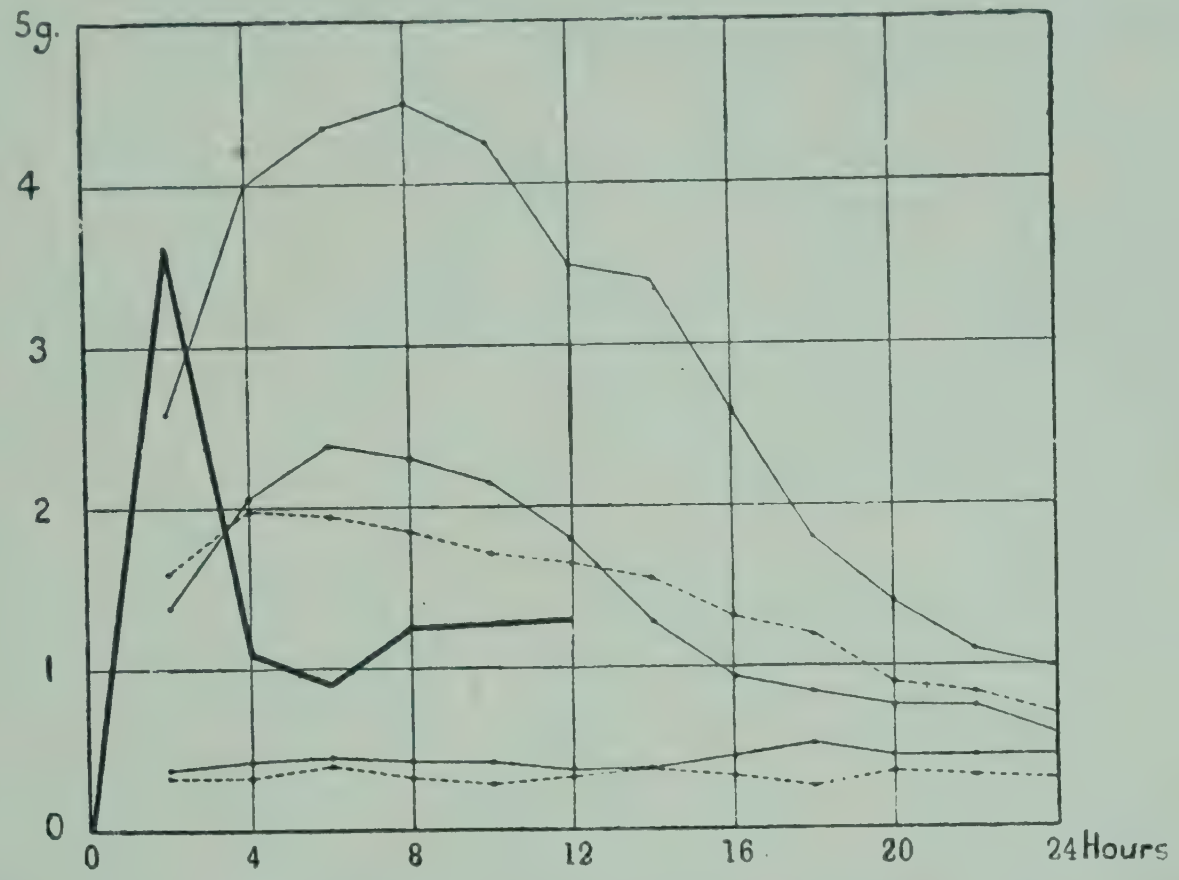


FIG. 587. The Hourly Variation in the Excretion of Nitrogen after a Meal.

The meal was given at 0. The thick line represents the average absorption of the food from the alimentary canal. The thin-lined curves, in order from above, represent the N excretion (1) after a meal of 1,000 grammes meat ; (2) after 500 grammes meat and 150 grammes fat ; (3) after a meal of 500 grammes meat ; (4) and (5) both represent the excretion in a fasting animal. (From TIGERSTEDT after FEDER.)

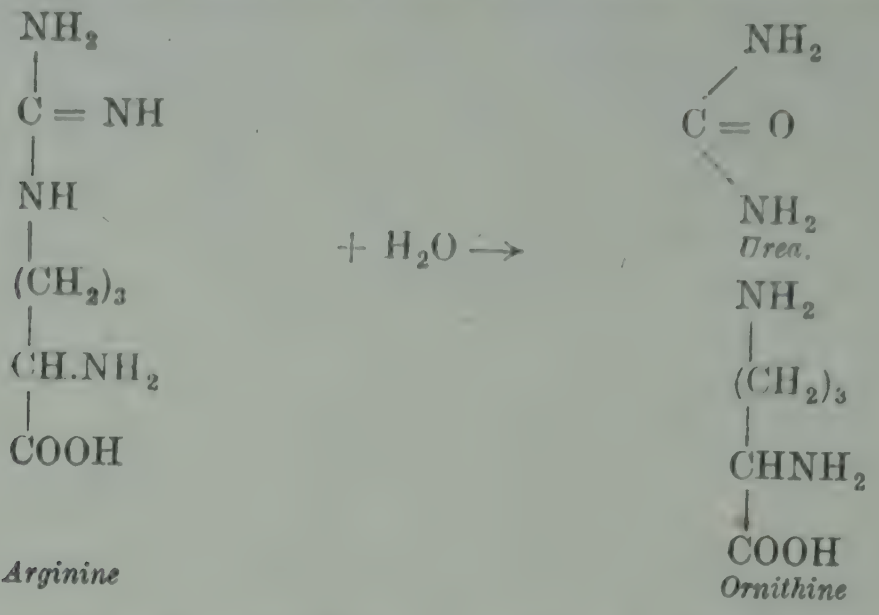
Urea Formation. It is evident that urea is the principal nitrogenous end-product of protein metabolism, both absolutely and relatively. On giving a large protein meal to a dog, the urea in the urine rapidly rises, and at the end of four or five hours 50 per cent. of the total nitrogen ingested with the food has appeared in the urine as urea (Fig. 587). Hence by far the greater portion of the protein nitrogen taken with the food is excreted as urea almost directly after absorption.* Urea is therefore to be regarded in the first place as an index to the amount of protein absorbed. That the nitrogen of the amino-acids is the immediate source of the urea is shown by the fact that the administration of these bodies is followed very rapidly by the appearance of the whole of their nitrogen in the urine as urea †: also, feeding with isotopic (N^{15}) amino-acids leads to excretion of isotopic urea.

It has long been known that the formation of urea from arginine is accomplished by the action of the enzyme *arginase*. It is abundant in th liver, and by its action arginine is split up into urea and ornithine *:

---

* PETTENKOFER and VOIT. *Ann. Chemie*, 1862, **52**, 361.

† LEVENE and MEYER. *Amer. J. Physiol.*, 1909, **25**, 214.



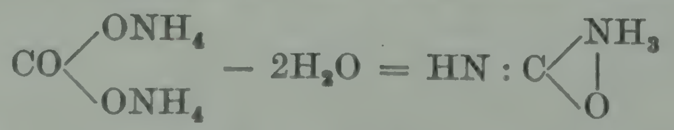
$$NH_2{-}C({=}NH){-}NH{-}(CH_2)_3{-}CH.NH_2{-}COOH + H_2O \longrightarrow NH_2{-}C{=}O{-}NH_2 + NH_2{-}(CH_2)_3{-}CHNH_2{-}COOH$$

In other instances the formation of urea is less simple than had been formerly supposed. The first stage is the formation of ammonia by *deamination* of the amino-acids, as previously described. In this change, oxygen is absorbed, and a keto-acid produced :—

$$R.CHNH_2.COOH + O \longrightarrow R.CO.COOH + NH_3$$

Tissue slice experiments confirm this, and show that liver and kidney are the tissues most active in deamination : the enzyme *l*-amino-acid oxidase, is a dehydrogenase. That ammonia can be a precursor of urea is shown by the fact that administration of isotopic NH_3 is followed by excretion of isotopic urea.

Although there is normally a small amount of ammonia in the urine, there is little if any in blood, and the urinary ammonia is not increased by injection or administration of ammonium carbonate which gives rise simply to a corresponding increase in the urea.† A large body of evidence points to the liver as being the chief seat both of deamination of amino-acids, and of conversion of ammonia into urea. Thus, if defibrinated blood be mixed with ammonium carbonate and passed for one hour through a surviving liver, the ammonium carbonate disappears and its place is taken by urea ‡ :—



$$CO\begin{cases}ONH_4\\ONH_4\end{cases} - 2H_2O = HN : C\begin{cases}NH_3\\O\end{cases}$$

though, as shown, below, the change is not so simple as this.

Extirpation of the liver in mammals is difficult, but by indirect procedures Mann § has succeeded in removing the liver from dogs.

After extirpation of the liver, both the deamination of the amino-acids and the formation of urea cease completely, indicating that under normal conditions both these processes are carried out in the liver. The animals only survive the completed operation for a few hours, but during that time the urea content of the blood and urine steadily falls. If the kidneys are also removed, the urea of the blood remains constant. In the normal animal, amino-acids are rapidly changed to urea, but when amino-acids are injected into the blood after hepatectomy, the greater part is recoverable from the blood and tissues, and there is no rise in urea or ammonia.

The actual steps by which ammonia is converted into urea, according to Krebs, depend upon the catalysing action of small amounts of ornithine, and on the presence of arginase.* In the liver, ornithine combines with ammonia and carbon dioxide to form first citrulline, and then with a molecule of ammonia to form arginine ; the arginine is then decomposed into urea and ornithine as explained above, and the ornithine thus set free can repeat the cycle, the reactions being :—

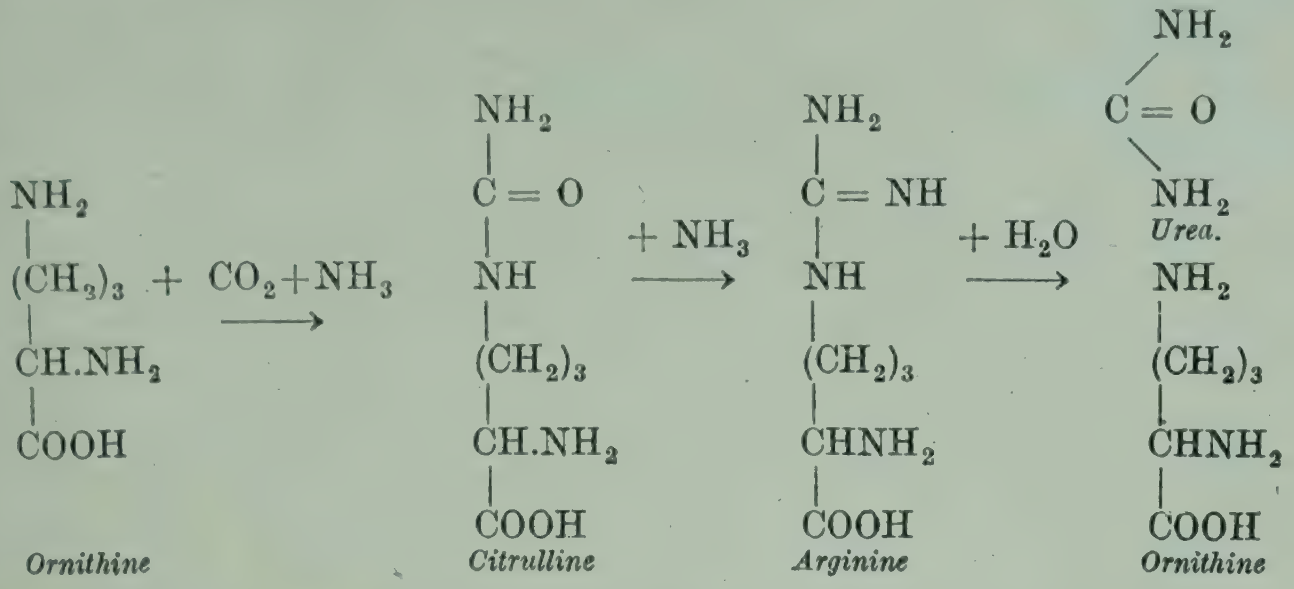
---

* Kossel and Dakin. *Z. f. physiol. Chem.*, 1904, **41**, 321.

† Knieriem. *Z. f. Biol.*, 1874, **10**, 263.

‡ Schröder. *Arch. f. exp. Path.*, 1882, **15**, 364; Löffler, *Bioch. Ztschr.*, 1918, **85**, 230.

§ Mann, *Amer. J. Physiol.*, 1921, **55**, 285; Mann and Magath, *Ergebn. d. Physiol.*, 1924, **23**, 212; Soskin, *J. Lab. Clin. Med.*, 1931, **16**, 382 (simplified).



The presence of oxygen, and some energy-yielding oxidative processes, *e.g.* oxidation of glucose or lactate, is essential, in order to provide the energy for the formation of urea.

There are possibly other paths by which urea can be formed.

Fate of the Amino-acids. The formation of urea does not signify a total disintegration of the amino-acid molecule, but is merely the result of its deamination.

The investigation of the disintegration of aliphatic derivatives generally, is rendered difficult by the fact that all the intermediate products undergo further change and leave the body in a state of complete oxidation as carbon dioxide and water. Much of our knowledge of these intermediate products is derived from the study of pathological conditions, in which they cannot be further oxidised, or by artificially converting the aliphatic intermediaries into derivatives, *e.g.* isotopic derivatives, or phenyl derivatives, which can be subsequently traced. If an amino-acid group be conjugated with an aromatic compound, *e.g.* forming a side chain of the benzene ring, it is protected from complete oxidation by this ring. The oxidation of the fatty side chain may only proceed to a certain degree, so that intermediate products of metabolism may be excreted still attached to the benzene nucleus.

In the α-amino acids the point where disintegration first occurs is the α-group. Deamination is most usually associated with oxidation.† The primary product is therefore an α-keto-acid. Thus from alanine the body would produce pyruvic acid :

$$\underset{\textbf{Alanine}}{CH_3.CHNH_2.COOH} + O = HN_3 + \underset{\textbf{Pyruvic acid}}{CH_3.CO.COOH}$$

This mode of deamination has been demonstrated to occur with slices of liver, or kidney, suspended in solutions containing various amino-acids and exposed to air,‡ keto-acids and ammonia being produced. Kidney tissue has the advantage that the ammonia is not converted to urea, as it is by liver. The enzyme concerned is an *l*-amino-acid oxidase. The keto-acid may afterwards undergo reduction to a hydroxy-acid, or oxidation to a lower fatty acid, but *in vitro* these changes can be inhibited by adding an arsenite.

The loss of nitrogen from the molecule diminishes little the energy value of the products derived from the amino-acids of the body. Thus, after deduction of the heat of combustion of urea, the heat of combustion of alanine is 349 Cal per g. molecule, while that of pyruvic acid is 279.

---

* KREBS and HENSELEIT. *Klin. Woch.*, 1932, **11**, 757, 1137.

† NEUBAUER, *Z. f. physiol. Chem.*, 1910, **67**, 230 ; **70**, 1, 326 ; KNOOP, *Z. f. physiol. Chem.*, 1910, **67**, 489 ; KNOOP and KERTESS, *ibid.*, 1911, **71**, 252.

‡ KREBS. *Z. f. physiol. Chem.*, 1933, **217**, 191 ; **218**, 157.

The important problem in the *energy* metabolism of protein is thus the nature and fate of the substances left after deamination. We have seen that protein exerts a high specific dynamic action (30 per cent.). This specific dynamic action of protein is possessed also by certain of the amino-acids. The action is not due to complete oxidation of the amino-acid, as is shown by giving glycine to a diabetic animal. The glycine is converted quantitatively into glucose, which is excreted in the urine, but yet produces the same augmentation of metabolism as in a normal animal. According to one theory, the specific dynamic action of protein is due to the amino-acids stimulating the cells of the body. But it cannot be that the amino-acids themselves are stimulants. This is shown by the fact that, when amino-acids are built up to form new tissues, as in the baby, or in the animal recovering from starvation, they exert no specific dynamic action. This action occurs only *when the amino-acids are deaminated and then undergo conversion to urea*, and, in proof of this, it remains in abeyance after removal of the liver. The interesting fact that ammonium salts also cause an increased heat production,* and are also converted to urea in the liver, lends further support to the opinion that the heat of specific dynamic action is set free in connection with this conversion. It probably represents the heat of concurrent oxidations which are coupled with the formation of urea.

The molecule which remains after removal of the amino group is, as we have shown, probably the corresponding α-ketonic acid : the subsequent history of this differs according to its nature. Ultimately it is fully oxidised to CO_2 and water, but in its intermediary changes it may be merged in the stream of the intermediaries of either carbohydrate or fat metabolism.

The amino-acids are divisible into three groups, according as they form, as intermediate compounds, sugar, aceto-acetic acid, or neither of these substances. Thus, of the amino-acids formed by the digestion of proteins, glycine, alanine, serine, cysteine, aspartic acid, glutamic acid, hydroxy-glutamic acid, ornithine, arginine, proline and hydroxy-proline can under appropriate circumstances be converted quantitatively into glucose. Probably those residues which ultimately form glucose are converted first into pyruvic acid, from which glucose is readily formed in the liver. Since the formation of glucose reduces the formation or appearance of ketone bodies, these amino-acids are classed as *antiketogenic* amino-acids. On the other hand, leucine, isoleucine, phenyl alanine and tyrosine yield no glucose, even in the diabetic animal ; but may in the liver undergo conversion into aceto-acetic acid, which is also a stage in the oxidative disintegration of fats. This group is therefore classed as *ketogenic* amino-acids. Tryptophan, valine, lysine and histidine do not form either product.

The Cyclic Amino-acids. The chief of these amino-acids are tyrosine, phenyl alanine and tryptophan, and we may assume that they participate in the labile equilibrium between the tissue proteins and the free amino-acids.

Ultimately they can all be completely destroyed, with the production of urea, carbon dioxide and water. In this respect they present a contrast to almost all other compounds of the aromatic series. Thus phenyl propionic acid, when administered, is passed in the urine as hippuric acid (benzoyl glycine). But the ketonic acids corresponding to tyrosine or phenyl alanine (*p*-hydroxy-phenyl-pyruvic acid or phenyl pyruvic acid) when given to phlorrhizinised animals yield aceto-acetic acid.

Tyrosine and phenyl alanine are not only constituents of many proteins, but are also starting points in the body for the formation of two important hormones, *adrenaline* and *thyroxine*. Adrenaline is possibly formed by the action of appropriate enzymes, tyramine being an intermediate stage. Thyroxine is probably formed, in the thyroid, from diiodotyrosine, which is formed from tyrosine. Tyrosine is also the starting point for the formation of *melanin*, the pigment of hair and skin, which is formed by the action of an oxidase called tyrosinase.

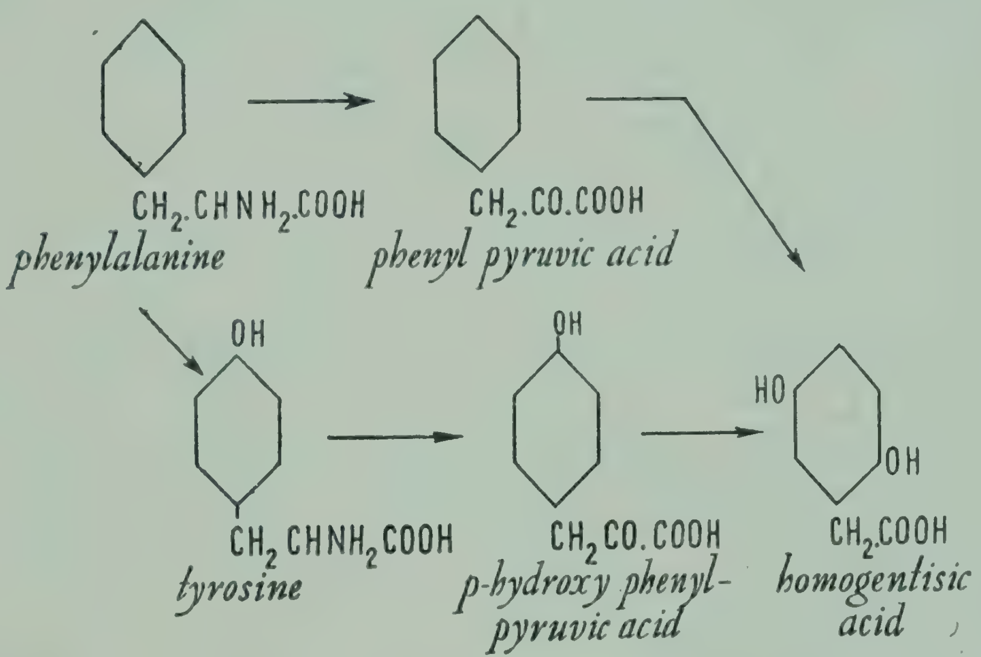
---

* LUNDSGAARD. *Skand. Arch. f. Physiol.*, 1931, **62**, **243**.

Phenyl alanine containing D when given by mouth yields tyrosine containing D in the liver, but the reverse change does not occur. Hence, phenyl alanine is essential, tyrosine is not.

The stages in the normal metabolism of tyrosine are not known. Some light is thrown on the subject by the study of a rare disorder in metabolism, known as *alcaptonuria*. In this condition, which is congenital, the urine darkens when made alkaline and exposed to the air, and reduces Fehling's solution. The peculiarities of the urine are due to the presence in it of a substance known as *homogentisic acid*, or 2,5-dihydroxyphenyl acetic acid.

If tyrosine or phenyl alanine be administered to an alcaptonuric it is quantitatively excreted as homogentisic acid. The constant ratio of this acid to the total nitrogen indicates that the whole of the tyrosine and phenyl alanine set free in the body is converted into homogentisic acid. It is probable that phenyl alanine and tyrosine are first oxidatively deaminised to phenyl pyruvic and *p*-hydroxyphenylpyruvic acid respectively, and these further oxidised to homogentisic acid :—



In a rare form of insanity, *imbecillitas phenylpyruvica*, phenylpyruvate oxidation is deficient and the acid is excreted. Phenyl pyruvic acid is also converted in alcaptonuric patients to homogentisic acid (instead of into aceto-acetic acid as it presumably is in the normal).

It is possible that under normal circumstances tyrosine and phenyl alanine are first oxidised to homogentisic acid, as in the alcaptonuric patient. Against this is the fact that *p*-hydroxyphenyl-pyruvic acid is converted into aceto-acetic acid in the liver. In the normal individual, however, the introduction of two hydroxyl groups into the benzene ring leads to an enzyme action which breaks up the ring. This enzyme is absent from the alcaptonuric, so that the transformation of the phenyl derivatives stops short at the stage of homogentisic acid.

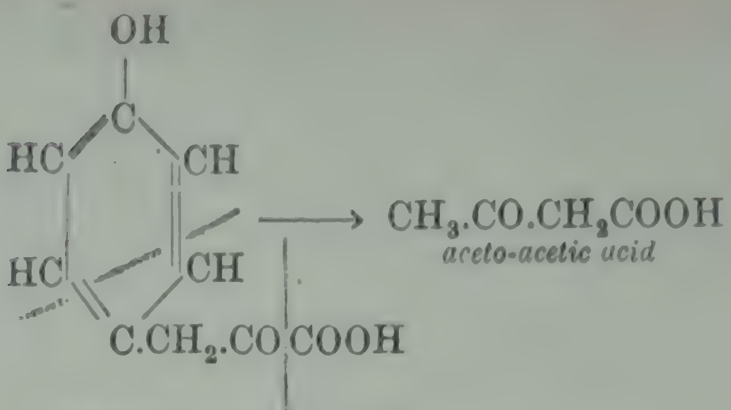
In the opinion of Dakin,* however, the formation of homogentisic acid from tyrosine is abnormal, and he considers that normally the benzene ring is split immediately after the formation of hydroxyphenyl-pyruvic acid, the side chain and parts of the ring forming aceto-acetic acid, while the rest of the ring is oxidised, as shown overleaf.

In another pathological condition (only once reported), the excreted product was *p*-hydroxyphenyl-pyruvic acid, and homogentisic acid if given was destroyed as it is by normal subjects. This supports Dakin's view.

Tyrosine in the intestine may undergo deamination by the action of bacteria, to form hydroxyphenyl-propionic acid and hydroxyphenyl-acetic acid. These

---

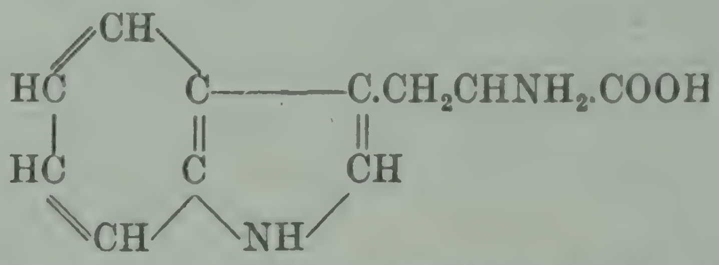
* DAKIN. *J. Biol. Chem.*, 1911, **9**, 151.



p-hydroxyphenyl-pyruvic acid

cannot be further oxidised, but appear in the urine as such or, after conversion into cresol or phenol, as sulphuric acid esters.

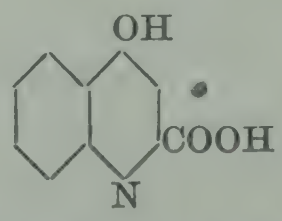
Tryptophan * is indol amino-propionic acid :—



It undergoes, like tyrosine, complete oxidation in the body, being probably first converted to the corresponding ketonic acid, indol-pyruvic acid, which can, therefore, replace it, in diets. The further stages are unknown.

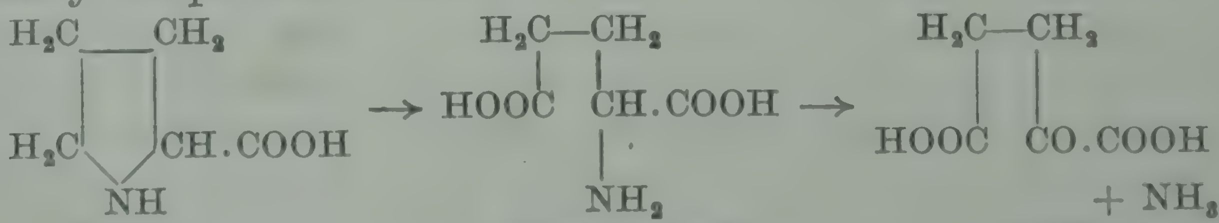
A very slight alteration in the molecule renders it incapable of this change. Thus some of the tryptophan may, under the influence of the bacteria of the intestine, be converted to indol acetic acid, and this by decarboxylation into skatol, or, by oxidation and further loss of carbon dioxide, into indol. Of these products, indol acetic acid may be found in the urine, and indol and skatol are oxidised to the corresponding phenols and pass into the urine conjugated either with sulphuric acid to form "urinary indican," or with glucuronic acid. The amount of the indol and skatol esters serves therefore as an index of bacterial decomposition in the alimentary canal. If putrefaction be prevented by an intestinal antiseptic, these esters may entirely disappear from the urine. On the other hand, a partial obstruction to the onward passage of food increases the indican excretion twenty or thirty-fold.

In the dog, tryptophan is converted into *kynurenic acid*, a quinoline derivative



which is excreted in the urine,† but this is probably not a stage in its oxidation in man.

THE PYRROL RING, which occurs in proteins as proline and hydroxyproline (*i.e.* pyrrolidine carboxylic acid and hydroxypyrrolidine carboxylic acid) appears to undergo complete disintegration in the body, by way of the tricarboxylic cycle.‡ Glutamic acid and α-ketoglutaric acid are probable intermediary compounds.



Proline *Glutamic acid* *α-ketoglutaric acid*

Both substances readily yield carbohydrate in the body.

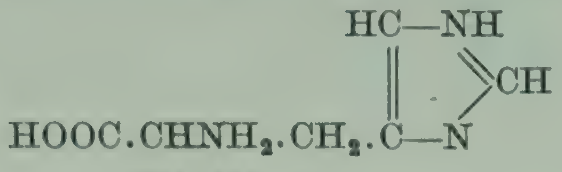
---

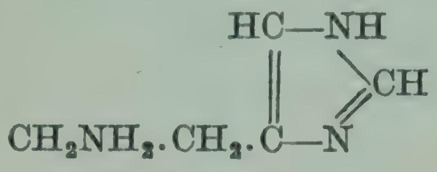
* KOTAKE. *Ergebn. d. Physiol.*, 1935, **37**, 245.

† ELLINGER, *Z. f. Physiol. Chem.*, 1904, **43**, 325; ELLINGER and MATSUOKA, *Z. f. Physiol. Chem.*, 1920, **109**, 259.

‡ TAGGART & KRAKAUR, *J. Biol. Chem.*, 1949, **177**, 641.

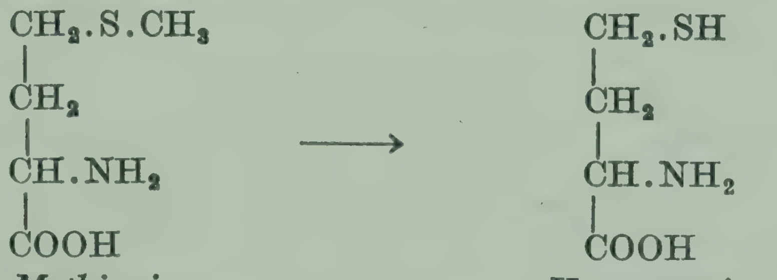
HISTIDINE is iminazol α-amino-propionic acid : ergothioneine and carnosine are derivatives of it. Histidine undergoes complete oxidation in the human body. The stages are unknown, but isotopic research suggests that the amino group of the alanine residue is very labile. The important derivative *histamine* is formed from it by the action of the enzyme histidine decarboxylase. In dogs, administration of histidine leads to excretion of urocanic acid in the urine. This is perhaps formed from histidine by the action of bacteria in the intestine.*


Histidine


Histamine

Amino-acids containing Sulphur. The normal source of sulphur is in the cystine and methionine links in proteins : it also exists in cysteine, glutathione and insulin, and in ergothioneine. Its ultimate fate is by excretion as the taurine of taurocholic acid in bile, by shedding of hair and nails (as keratin) and by excretion in the urine as sulphates, etc.

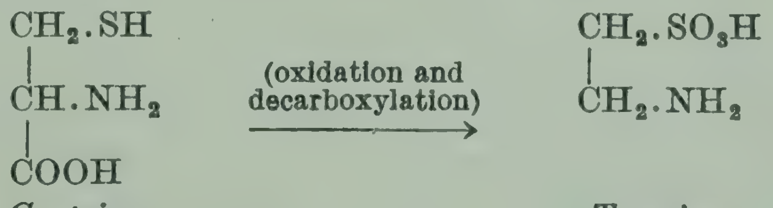
Methionine must be supplied to the body : the other ones need not be, and can be made from methionine, but we do not know by what changes. Methionine readily yields up its methyl group, and is converted into homocysteine : in the body this probably often happens, the CH_3 group being introduced into another molecule (*transmethylation*), and the homocysteine being then converted into cysteine by way of an intermediary reaction with serine. It is uncertain whether methionine can be converted into keto-bodies or sugar except *viâ* cysteine.



$$\begin{array}{l} CH_2.S.CH_3 \\ | \\ CH_2 \\ | \\ CH.NH_2 \\ | \\ COOH \end{array} \longrightarrow \begin{array}{l} CH_2.SH \\ | \\ CH_2 \\ | \\ CH.NH_2 \\ | \\ COOH \end{array}$$

Methionine *Homocysteine*

Cystine and cysteine are glycogenic. In the course of metabolism, the SH group is easily oxidised to a sulphonic group, *e.g.* :



$$\begin{array}{l} CH_2.SH \\ | \\ CH.NH_2 \\ | \\ COOH \end{array} \xrightarrow[\text{decarboxylation)}]{\text{(oxidation and}} \begin{array}{l} CH_2.SO_3H \\ | \\ CH_2.NH_2 \end{array}$$

Cysteine *Taurine*

From such compounds, sulphuric acid would most probably be set free.

Certain individuals have an inborn error of metabolism, and excrete relatively large amounts of cystine in the urine. Feeding them with cystine does not increase the amount secreted, but increases the sulphur excretion as in normal subjects : the explanation of this anomaly has not been found.

Sulphur occurs in the urine in three forms : as inorganic sulphates, as ethereal sulphates (indoxyl- and skatoxyl-sulphates), and in an unoxidised condition often termed neutral sulphur. This last consists of cystine, thiocyanates, and, in some animals, mercaptan compounds. The excretion of sulphur rises *pari passu* with that of the urea so that, with the throwing off of the NH_2 group, there must be also a removal of the greater part of the sulphur of the protein molecule. The sulphur does not necessarily appear in the urine simultaneously with the nitrogen derived from the same protein. If, for instance, there is a positive N-balance, *i.e.* if protein is being built up more rapidly than broken down, the excretion of S lags behind, and in the reverse condition of excessive

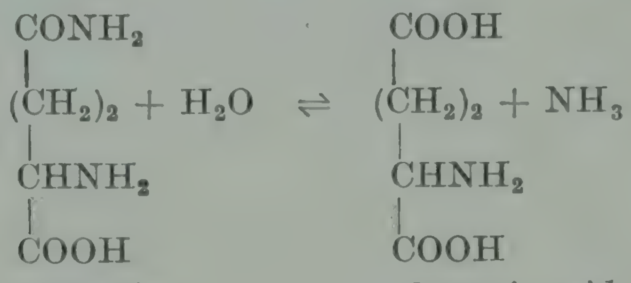
---

* RAISTRICK. *Bioch. J.*, 1916, **11**, 71.

N-catabolism the N lags behind. On the whole, however, the N : S excretion ratio is the same as that of the protein consumed, or, in starvation, as that of muscle proteins, viz., about 14.

The ethereal sulphates are excreted in varying quantity, according to the extent of the decomposition processes which are occurring in the intestine, whereby tryptophan is converted into indol and skatol, which are then detoxified by conversion to their sulphates, as already explained. On the other hand, the fact that the neutral sulphur undergoes no decrease on starvation suggests that this part of the sulphur output may be connected with tissue metabolism. Neutral sulphur excretion amounts to about 100 mg. daily, but varies widely ; about 8 per cent. of it is in the form of cystine, of which an average of about 30 mg. is excreted daily.*

The Excretion of Ammonia. A small proportion of the nitrogen in the urine occurs in the form of ammonia. The ammonia of the urine is formed not in the liver, but in the kidneys, and is not derived from ammonia in the plasma. Hence, feeding with isotopic ammonia or urea, leads to the excretion of isotopic urea, but no isotopic ammonia. Its precursor is glutamine of the plasma, and the kidney contains the enzyme glutaminase, which effects the conversion :—



$$\begin{array}{lcl} \mathrm{CONH_2} & & \mathrm{COOH} \\ | & & | \\ \mathrm{(CH_2)_2} + \mathrm{H_2O} & \rightleftharpoons & \mathrm{(CH_2)_2} + \mathrm{NH_3} \\ | & & | \\ \mathrm{CHNH_2} & & \mathrm{CHNH_2} \\ | & & | \\ \mathrm{COOH} & & \mathrm{COOH} \end{array}$$

glutamine *glutamic acid*

By this reversible reaction, ammonia, which is a very toxic substance, can be stored or released as required. The factor which determines the proportion of ammonia in the urine is the relative proportion of acid and base, i.e., of $H^{\cdot}$ and OH' ions, which has to be eliminated from the body, and this also determines the ratio of dibasic to monobasic phosphates excreted. If the fixed alkalies in the food are sufficient to combine with the whole of the acids excreted from the body, then the urine is neutral and ammonia will be absent. If, however, a dose of mineral acid be administered to an animal, this must be excreted in combination with some of the base ; the urine will then be acid, owing to a preponderance of monobasic phosphate, but some of the acid will be excreted in combination with ammonia. The ammonia of the urine is therefore an index to the amount of acids excreted. These acids may be introduced directly with the food, as when mineral acids are administered by the mouth, or may be the product of metabolic processes. Thus, if ammonium chloride be given to a starving rabbit much of it is *apparently* excreted ; actually, the ammonia is converted to urea, leaving HCl behind, and new ammonia is made in the kidney as a result. In carbohydrate starvation, hydroxy-butyric acid and aceto-acetic acid are produced in the body in large quantities, but cannot undergo further disintegration. It is therefore essential that the acids formed should be neutralised and excreted in combination with some base. When the alkalies of the food and of the tissues do not suffice for their neutralisation, ammonia is formed in the kidneys and is then utilised for this purpose, and the acids appear in the urine in combination with ammonia. The ammonia of the urine therefore gives valuable information as to the formation of acids in abnormal quantities during the processes of metabolism (" acidosis ").

For Reference

BALDWIN (1936). Arginase. *Biol. Rev.*, **11**, 247.
BOLLMAN and MANN (1936). The Physiology of the Impaired Liver. *Ergebn. d. Physiol.* **38**, 445.

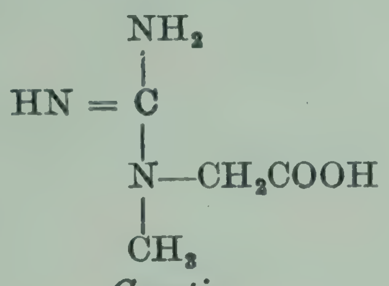
---

* MEDES. *Bioch. J.*, 1937, **31**, 12.

Du Vigneaud (1942). The Significance of Labile Methyl Groups in the Diet and Their Relation to Transmethylation. (*Harvey Lects.*, 1942–43, p. 39.)
Fearon (1926). The Biochemistry of Urea. *Physiol. Rev.*, **6**, 399.
Lewis (1924). Sulphur Metabolism. *Physiol. Rev.*, **4**, 394.
Lewis (1941). The Significance of the Sulphur-containing Amino-acids in Metabolism. (*Harvey Lects.*, p. 159, Lancaster, Pa.; Science Press Printing Co.)
Mann and Magath (1924). Die Wirkungen der totalen Leberextirpation. *Ergebn. d. Physiol.*, **23**, 212.
McCance (1930). The Chemistry of the Degradation of Protein Nitrogen. *Physiol. Rev.*, **10**, 1.

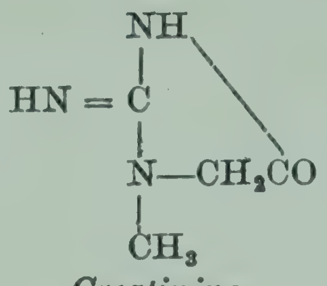
Creatinine and Creatine. On comparing the output of the various nitrogenous excreta given in Folin's Table quoted above (p. 924), we see that on a low protein diet the only substance which does not undergo simultaneous diminution is the creatinine. At one time this led to the belief that creatinine might be regarded as an index of tissue metabolism of protein.

Creatinine is derived by dehydration from creatine (methyl guanido-acetic acid).



$$HN=C(NH_2)-N(CH_3)-CH_2COOH$$

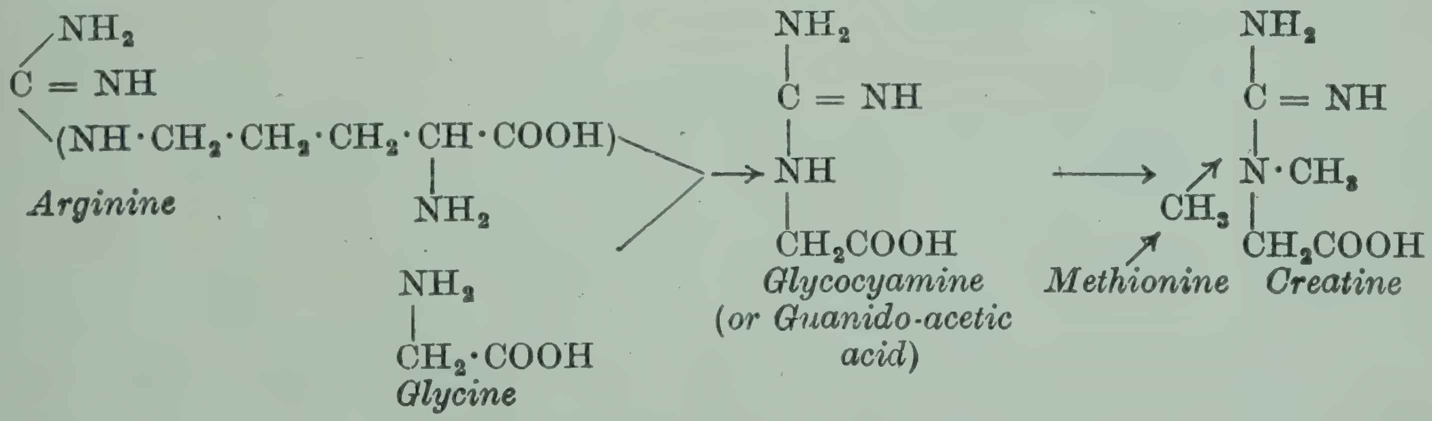
Creatine



Creatinine

Both are said to be present in the blood, creatine to the extent of about 3–5 mg./100 c.c., mostly present in the red cells, and "creatinine" about 2 mg./100 c.c. Most of the apparent creatinine of blood, however, is some other, unidentified, substance. Muscle contains about 10 mg. creatinine and about 200 mg. creatine (as creatine phosphate) per 100 grammes, so that the body of a man contains about 100 g. of this substance.

How or where creatine is formed in the body is uncertain, though its most probable site of formation is in the muscles and perhaps in the liver. By the use of compounds containing isotopic nitrogen it has been shown that, in the formation of creatine, arginine, glycine, and probably methionine play parts, the last providing CH_3 :—



In progressive muscular dystrophy, in which creatine is always present in the urine, the creatine excretion is increased by the administration of either glycine or arginine.* Creatine is also excreted by many normal persons after administration of glycine.†

Creatinine is certainly to be regarded as a waste product. About 1–1·5 grammes per day is excreted. When given *per os* it is excreted in the urine. Since isotopic creatine given *per os* leads to excretion of isotopic creatinine, it seems certain that the urinary creatinine is derived from creatine. But creatinine is excreted in considerable quantities even when a creatine-free diet is given. It has been found, moreover, that creatine administered in moderate amount by the mouth disappears in the body. It would seem probable that this is due to its conversion into creatine phosphate, which is stored in the muscles, of which it forms an invariable and essential constituent. If a large amount be given, creatine appears as such in the urine. In most cases also a certain small proportion is excreted in the form of creatinine.* After repeated administration of creatine by mouth, steadily increasing amounts of creatine and creatinine are excreted in the urine.† This would make it appear that excretion as creatine only occurs when the creatine store of the body exceeds a certain limit : it may be stored for a long time in the muscle. Hence, if isotopic creatine is given for a considerable time, and followed by a creatine-free diet, the excretion of isotopic creatinine continues.

---

* Brand *et al.* *Amer. J. Physiol.*, 1929, **90**, 296; *J. Biol. Chem.*, 1930, **87**, ix.; 1932, **92**, lix.; Beard and Barnes, *J. Biol. Chem.*, 1932, **94**, 49.

† Adams, Power and Boothby. *Amer. J. Physiol.*, 1935, **111**, 596.

Creatine is normally present in the urine of children before puberty,‡ and of women, especially in pregnancy, but not in that of normal adult men, except when they are starved, or deprived of carbohydrate. It also appears in hyperthyroidism and in various diseases, *e.g.* progressive muscular dystrophy and myasthenia gravis associated with muscular atrophy or weakness. Probably in most cases the excretion is due to failure of storage.

With creatinine it is different : once this substance is introduced into the body it is promptly excreted. The amount formed certainly bears some relation to the amount of muscle present, though not to the amount of muscular work done, when long periods are considered. Thus a muscular man may excrete up to 30 mg. creatinine per kg. body-weight per day (creatinine coefficient), whereas in women it is usually about 18 mg. and in children about 15 mg.

The excretion of creatinine was investigated under various conditions by van Hoogenhuyze and Verploegh, and supports the view expressed above as to the intimate relation of creatinine with the tissue metabolism. During and immediately after muscular exercise, there is an accelerated output of creatinine in the urine, but there is a corresponding drop during recovery, so that the daily excretion is but little affected. Since we know that during muscular contraction there is a conversion of creatine phosphate into creatine and inorganic phosphate, both of which are diffusible, it is reasonable to suppose that some of the creatine escapes into the general circulation, is converted into creatinine, and excreted as such. During recovery any available creatine would be taken up again, as creatine phosphate is re-formed.

For Reference

Hunter (1922). The Physiology of Creatine and Creatinine. *Physiol. Rev.*, **2,** 586.
Rose (1935). The Metabolism of Creatine and Creatinine. *Ann. Rev. Biochem.*, **4,** 243.

2. NUCLEIN METABOLISM

An outline of the chemistry of nucleoproteins has already been given in Chapter IV.

The nucleoproteins are especially abundant constituents of nuclei, and therefore occur to a greater or less extent in all the ordinary animal foods, eggs and milk excepted. Other important purine derivatives are the mononucleotides, adenylic acid and guanylic acid, which occur in the cytoplasm of many tissue cells, *e.g.* in muscle, where adenylic acid is present as adenosine triphosphate and plays an important part in its metabolism. The metabolism of the nucleoproteins and nucleotides is essentially comprised in the history of their characteristic constituents, the purines and pyrimidines. The protein portion of the molecule, which contains a high proportion of diamino acids, shares the common fate of other protein.

---

* Rose and Dimmitt. *J. Biol. Chem.*, 1916, **26,** 345.
† Benedict and Osterberg. *J. Biol. Chem.*, 1923, **56,** 229.
‡ Rose. *J. Biol. Chem.*, 1911, **10,** 265.

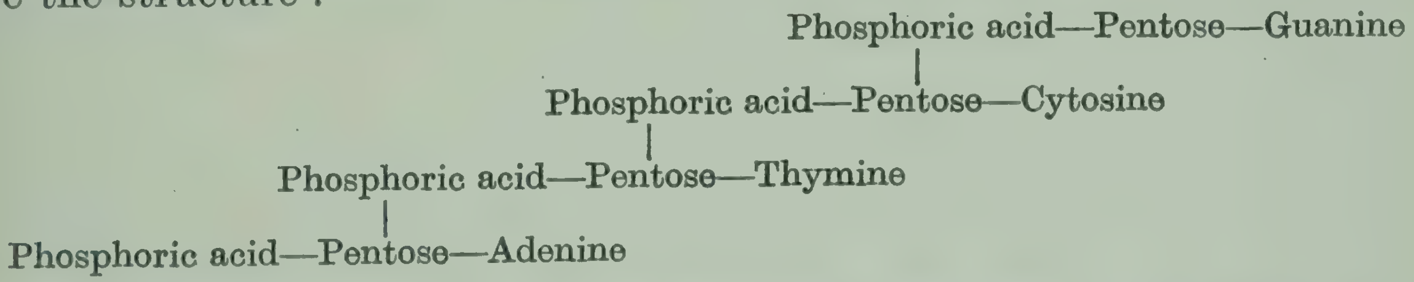
Formation of Nucleins in the Body. We have seen that, in the higher animals proteins are built up from their constituent amino-acids, some of which can be synthesised in the body. The nucleotides also can certainly be synthesised by the animal. This is shown by the fact that the hen's egg before incubation contains practically no nuclein or purine bases. During incubation there is a rapid increase in the number of nuclei, so that the chick just before it is hatched contains a considerable amount of nuclein. These must have been formed by synthesis from the phosphoproteins and phosphatides which form so important a constituent of the egg-yolk.* In the same way, in the young infant there is a rapid increase in the nucleoprotein of the body, although the only food ingested is milk, which contains but an insignificant amount of purine compounds. Confirmation of the synthesis of purines in the body has also been given by Benedict: he fed a Dalmatian hound on a purine-free diet, yet was able to extract 100 grammes of uric acid from its urine.†

Experiments with isotopic N-compounds show that the purines and pyrimidines are formed, not from amino-acids, but from ammonia, and indicates the site of formation to be the liver, and perhaps other sites, such as kidney.‡

Fate of Nucleoproteins in the Body. Nucleoprotein metabolism may be exogenous or endogenous, *i.e.* the nucleoprotein broken down in the body may be derived either from the food or from the tissues. Most of our knowledge on this subject relates to the purine portions of the molecule, the pyrimidine part having been much less explored. We will first deal with the exogenous nucleoprotein.

In the alimentary canal, nucleoproteins are acted upon by the pepsin of the gastric juice and trypsin of pancreatic juice, the protein being digested away and nucleic acid remaining.

The nucleic acids are compounds of very high molecular weight, formed by aggregation of molecules of tetranucleotides, in which four mononucleotides are linked together. Thus, a tetranucleotide of animal origin would probably have the structure :—


Phosphoric acid—Pentose—Guanine
Phosphoric acid—Pentose—Cytosine
Phosphoric acid—Pentose—Thymine
Phosphoric acid—Pentose—Adenine

The intestinal mucosa contains an enzyme, *nuclease*, which splits up the nucleic acid into its four constituent mononucleotides,§ *e.g.*

Phosphoric acid—Pentose—Adenine (= *Adenylic acid*)

is a purine nucleotide, and

Phosphoric acid—Pentose—Cytosine

is a pyrimidine nucleotide ; it is likely that some at least of these products are absorbed as such, since an adenine nucleotide has been recovered from human blood.‖ However, further changes can be effected by the intestinal juice, or by the cells of the intestinal mucosa. The latter contain the enzyme *nucleotidase*, a phosphatase, ¶ by which the phosphoric acid is split off from the nucleotides, with formation of nucleosides, *e.g.*

Phosphoric acid—Pentose—Adenine ——→ Pentose—Adenine (= *Adenosine*, a nucleoside).

---

* Plimmer and Scott. *J. Physiol.*, 1909, **38**, 247.
† Benedict. *J. Lab. Clin. Med.*, 1916, **2**, 1.
‡ Barnes and Schoenheimer. *J. Biol. Chem.*, 1943, **151**, 123.
§ Levene and Dillon. *J. Biol. Chem.*, 1932, **96**, 461.
‖ Jackson. *J. Biol. Chem.*, 1924, **59**, 529.
¶ Levene and Dillon. *J. Biol. Chem.*, 1930, **88**, 753.

The nucleosides are further acted upon by another enzyme, also present i the intestinal mucosa, and called *nucleosidase* *; this breaks up the purine nucleosides into purine base and carbohydrate, *e.g.* :—

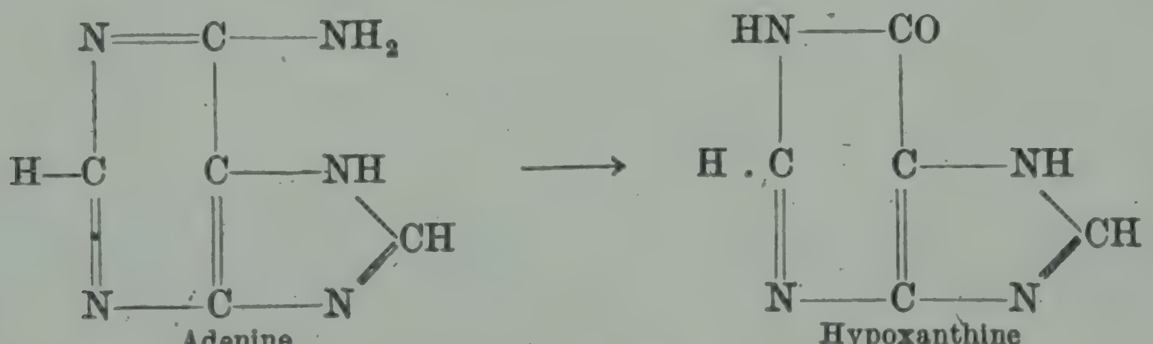
$$\begin{tikzcd} & \text{Adenosine} \arrow[dl] \arrow[dr] & \\ \text{Carbohydrate (Pentose)} & & \text{Adenine} \end{tikzcd}$$

There may therefore be present for absorption in the intestine, nucleotides, nucleosides, purine bases (guanine and adenine), carbohydrate and phosphoric acid. These products presumably may all enter the blood and be dealt with, together with similar products formed, we know not where or how, in endogenous purine metabolism, which we must now discuss.

The chief ultimate purine end product excreted in man is uric acid. In some animals it is allantoin. Ingestion of nucleoprotein, nucleic acid or purine bases is, in man, followed by an increased excretion of uric acid in the urine. It is evident that the uric acid of the urine may be derived either from the nucleins, etc., of the food, or from those of the tissues. By digestion of nucleic acids with animal tissues or their extracts under varying conditions, it is possible to bring about all the changes mentioned above, and also those further ones involved in the conversion of the purine bases contained in them into uric acid. Probably by no means all the purine bodies involved in metabolic changes are excreted as such, a good deal being lost in some unknown way.

Tissue-nucleotidases are of different kinds, phosphonucleotidases splitting off the phosphoric acid and leaving the nucleosides, while the purine nucleotidases split off the purines, leaving the phosphoric acid combined with the carbohydrate. (This could then be split into its components by a phosphatase.)

The purines set free in these various ways undergo further changes. One of these is deamination. Thus, guanine and adenine, acted upon by the enzymes guanase and adenase, are converted respectively into xanthine and hypoxanthine by loss of their amino groups, *e.g.*


Adenine
Hypoxanthine

But deamination can also occur before the purines are liberated, *i.e.* in the nucleosides. Thus adenosine can yield inosine by the action of an appropriate deaminase. By a similar change even the nucleotide can be deaminated, *e.g.* adenylic acid can be converted to inosinic acid, and from this a purine nucleotidase can split off hypoxanthine.

The hypoxanthine is converted under the action of an oxidase (xanthine oxidase) first into xanthine and then into uric acid. This was one of the earliest facts discovered in the metabolism of purines. If spleen pulp be digested with blood, it is possible to extract xanthine from the mixture.† If, however, oxygen be bubbled through the fluid, the xanthine disappears, its place being taken by uric acid. All these changes occur in the living body, though not necessarily in the order just set out. Thus a pulp of dog's liver is able completely to deaminise the adenine group of nucleic acid, converting it into a hypoxanthine group; but it is without effect on free adenine. It is evident, therefore, that the various enzymes which have been described act partly on the whole nucleotide molecule, partly on the products of its decomposition, and that the results of the action of the body enzymes are not the same in the two cases.

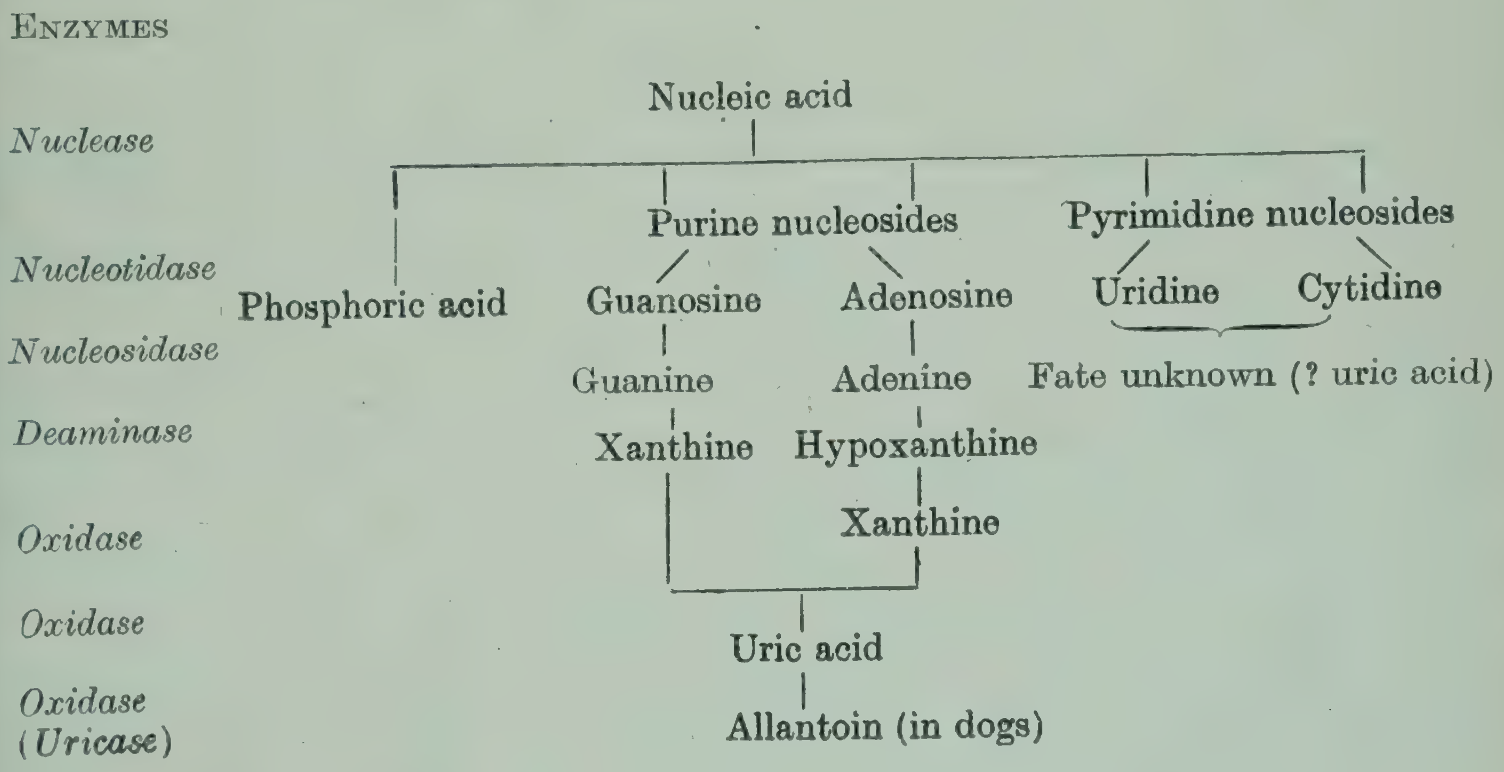
---

* Levene and Dmochowski. *J. Biol. Chem.*, 1931, **93**, 563.

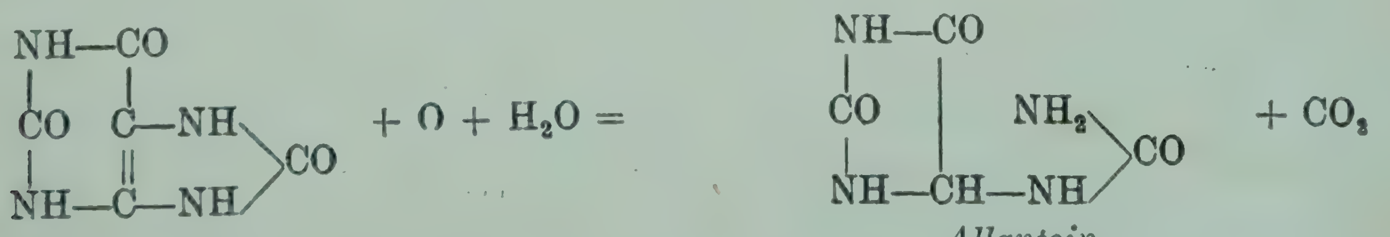
† Horbaczewski. *Monat. f. Chem.*, 1891, **12**, 221.

These various enzymes are not all present in all tissues, nor are the same ones present in the same tissues of different animals. Thus human liver contains guanase and xanthine oxidase; guanase is present in one or other tissue of most animals, except in those of the pig. Adenase is absent from human tissues *; guanase is present in human kidney, liver and lung, but not in spleen or pancreas. Xanthine oxidase is present in the human liver but not in other human tissues. For these reasons there are species differences in the way in which the purines are dealt with; for instance, the muscles of the pig often contain deposits of guanine because this cannot be destroyed, while adenine is present in human urine in small amounts. Yet not all the adenine present in the compounds ingested is excreted in man, because much of this purine is deaminised, by enzymes other than adenase, while still a part of the nucleotide or nucleoside molecule.†

If we make this reservation that the constituent parts of the nucleic acid molecule may undergo changes while still bound to the other parts, we may represent diagrammatically the formation of uric acid from nucleic acid as follows:—


ENZYMES
Nuclease
Nucleotidase
Nucleosidase
Deaminase
Oxidase
Oxidase
Oxidase
(Uricase)
Nucleic acid
Phosphoric acid
Purine nucleosides
Pyrimidine nucleosides
Guanosine
Adenosine
Uridine
Cytidine
Guanine
Adenine
Fate unknown (? uric acid)
Xanthine
Hypoxanthine
Xanthine
Uric acid
Allantoin (in dogs)

The question arises whether the uric acid excreted by a man represents the whole of the purines which have been destroyed in the body. Although complete equivalence has been found between the amount of hypoxanthine ingested and the amount of uric acid excreted, the same equivalence has not been established in the case of nucleic acid; and the important question arises whether uric acid once formed is stable or whether it may undergo further changes before being excreted. In many animals, such as the dog, the amount of uric acid in the urine is only minute, the chief purine derivative in this fluid being allantoin—an oxidation product of uric acid:


NH—CO
CO
C—NH
CO
NH—C—NH
+ O + H₂O =
NH—CO
CO
NH₂
CO
NH—CH—NH
+ CO₂

Allantoin

---

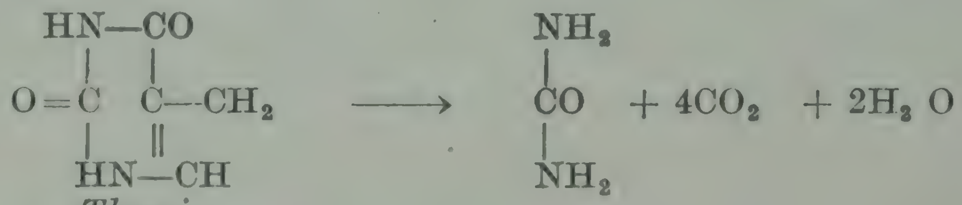
* JONES. *J. Biol. Chem.*, 1911, **9**, 129, 169.

† JONES and WINTERNITZ. *Z. f. Physiol. Chem.*, 1905, **44**, 1; 1909, **60**, 180.

This transformation can be effected by extracts made from the liver of the dog and various other animals.* The enyzme carrying out this change is known as uricase. No such enzyme is found in any human organs.†

If we add together the amounts of allantoin and uric acid excreted, and express the allantoin as a percentage of this sum, we get a figure called the *uricolytic index*. It is a measure of the extent to which uric acid is converted in the body into allantoin (by the enzyme uricase). In dogs of most species it is 98, *i.e.* most of the uric acid is oxidised to allantoin before excretion, but in the Dalmatian hound it is only 32. In the chimpanzee it is 0, and in man only 2. In man and the higher apes, therefore, practically all the uric acid formed in the body would appear to be excreted as such in the urine. The small trace of allantoin which may occur in human urine is directly derived from the food.

So far we have not considered the fate of the pyrimidine portion of the nucleotide molecule. It would seem that the human body is able completely to destroy these compounds, since they are absent from the urine. If a large amount of thymine or uracil is administered to dogs, some of it passes out in the urine unchanged, though, if given slowly it can be destroyed so that none is recoverable from the urine. Since in this condition the output of urea is increased, it would appear that the pyrimidines are converted into urea. Though we have no knowledge of the site or mode of this transformation, it is, chemically speaking, a simple one, as shown :



Thymine

Excretion of Uric Acid. Normal human blood contains 2 to 3 mg. of uric acid per 100 c.c., but the bloods of animals which do not excrete it contain very small amounts. When given intravenously, as lithium urate, to man, a variable fraction is excreted during the next few days ; the kidney does not seem able to deal with it at all rapidly. Under ordinary circumstances a man excretes about half a gramme of uric acid per day. In addition the urine contains a small amount of purine bases. From 10,000 litres of human urine Krüger and Salomon ‡ succeeded in isolating xanthine (10·1 grams), hypoxanthine (8·5 grams), adenine (3·5 grams). These are probably derived from the uric acid metabolism. In addition there are various methylated purines, derived from such constituents of the diet as caffeine, etc.

An attempt has been made to arrive at the amount of uric acid produced endogenously from a study of the quantity of uric acid in the urine of subjects on different diets. During starvation, when a man is living on his own tissues, one might expect the uric acid to be increased in consequence of disintegration of the tissues. It has been suggested that the amount of the endogenous uric acid in the urine would be obtained by an analysis of the urine from subjects taking a diet free from purine bases, but containing sufficient nitrogen to maintain nitrogenous equilibrium. Even in these circumstances, however, the amount of uric acid increases with the total nitrogenous metabolism, and, in fact, increases after each meal (Fig. 588).

There are three conditions which give rise to a definite increase in the output of endogenous uric acid. These are (1) severe muscular activity, (2) febrile states accompanied by increased nitrogenous metabolism. In both these conditions we may regard the uric acid as derived especially from the adenylic acid contained in the muscle. (3) In maturation of red cells, as, *e.g.* when a pernicious anæmia subject has an induced reticulocyte response resulting from the administration of liver extract. Presumably the increased uric acid is derived from the destroyed nuclei of the erythroblasts.

The foods which are especially effective in causing increase in the exogenous

---

* Wiechowski u. Wiener. *Hofmeister's Beitr.*, 1907, **9**, 247, 295.

† Jones. *Z. f. Physiol. Chem.*, 1909, **60**, 180.

‡ Krüger and Salomon. *Z. f. Physiol. Chem.*, 1898, **24**, 364 ; **26**, 350.

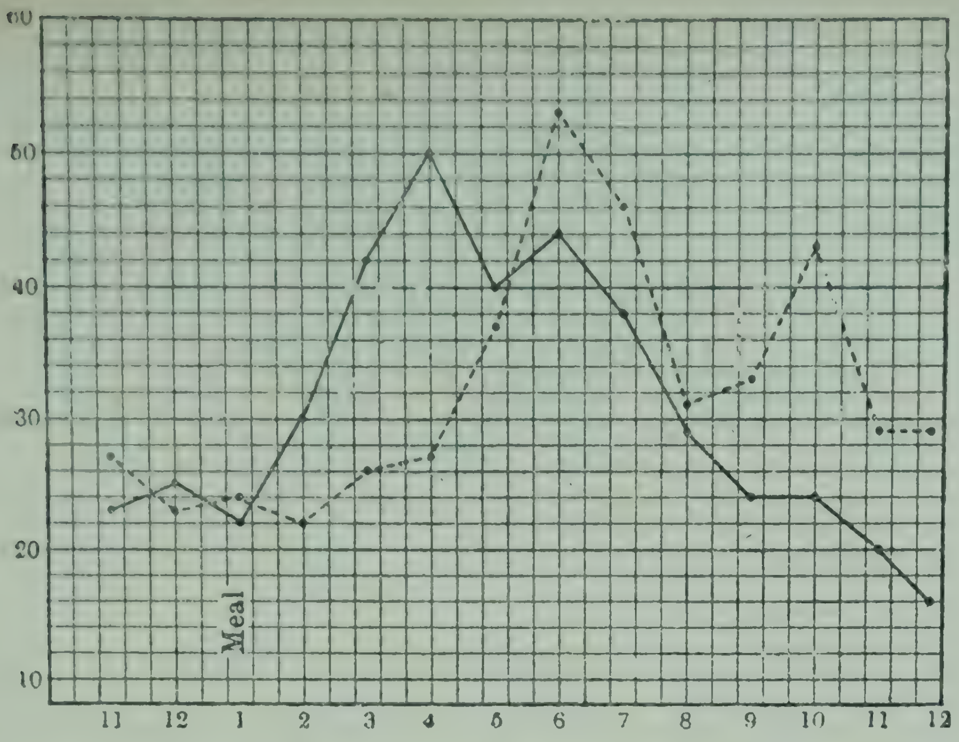


Fig. 588. Curves showing the Hourly Excretion of Uric Acid and Urea after a Single Meal. (Hopkins and Hope. *J. Physiol.*, 1898, **23**, 271.)

The continuous line = uric acid output; the dotted line = urea output.

uric acid are those rich in nuclein, such as sweetbreads or liver, and those rich in hypoxanthine or its precursors, such as meat or meat extract.

For Reference

Bredereck (1938). Nucleasen. *Ergebnisse der Enzymforschung*, **7**, 105.
Drury (1936). The Physiological Activity of Nucleic Acid and Its Derivatives. *Physiol. Rev.*, **16**, 292.
Jones (1920). Nucleic Acids, etc. (London: Longmans Green.)
Rose (1923). Purine Metabolism. *Physiol. Rev.*, **3**, 544.

3. FAT METABOLISM

The study of fat metabolism entails that of the metabolism of all compounds which contain fatty acids, and we shall begin therefore with a consideration of depôt fat.

Depôt Fat. We have seen that the amount of fat in the body is the main factor governing the length of life during starvation, because it is then the only material which can serve as a source of energy for a sufficient length of time. This reserve material, which is chiefly neutral fat, is found widely distributed in the body and in greatly varying amounts; the body of the average person contains some 12 per cent. of its total weight as fat, nearly all of which can act as an energy reserve. Its distribution in the body varies considerably, but we may say that the *panniculus adiposus* under the skin may account for, roughly, 50 per cent., and the perirenal, mesenteric and omental fat for about 15 per cent., 10 and 5 per cent. respectively, while some 5 per cent. or more may occur in the intermuscular connective tissue. In this adipose tissue the fat is enclosed within distended connective tissue cells, the protoplasm of which is reduced to a thin pellicle round each fat globule.

Composition of Depôt Fat. This reserve fat occurs in the form of triglycerides of various fatty acids the chief of which are stearic, palmitic and oleic acids. Ordinary depôt fat contains 75 per cent. of oleic acid, with 20 per cent. of palmitic acid and 5 per cent. of stearic acid, combined as mixed triglycerides, although often small amounts of more unsaturated acids also occur. There are also present small amounts of lecithin and cholesterol, which we may ignore for the moment.

The degree of saturation of depôt fat depends on its position in the body, there being a decreasing iodine value from the surface inwards. Thus, in a pig fed on barley the iodine values of the subcutaneous fat at increasing distances from the surface were 60, 57, 52 and 51 respectively; this is to be correlated with the increase of temperature as the distance from the skin increases,[*] and depends in part on the fact that the more unsaturated fats are liquid at lower temperatures than are the saturated ones. The same phenomenon is shown in the fact that the subcutaneous fat of the new-born infant has an iodine value of 47, which increases during the first year to the adult figure of 70.[†] This again is to be correlated with the internal position of the fœtus and the gradual adaptation during the first year of life to the fall in temperature of the subcutaneous tissues.

The Origin of Depôt Fat. (*a*) *The Food Fat.* It is a familiar experience that the fats of different species of animals usually have characteristic appearances, *e.g.* the hard fat of mutton, compared with the soft fat of beef. This relative constancy of composition within a species depends to a considerable extent on the similarity of the food of its members. But food fat may also undergo some modification during assimilation, which again will tend to make the fat of a particular animal of a constant nature. Although the nature of the depôt fat is capable of modification by the provision in the food of fats unusual in amount or nature, profound changes in its nature are usually only observed in experimental conditions so severe as to be unphysiological. The most graphic method of demonstrating such changes is to starve an animal until its fat reserves have been largely exhausted and then to fatten it on a particular fat; it is then found that the depôt fat approximates closely in composition to that of the fat given. In experiments of this type linseed oil caused in dogs an adipose tissue containing fat which was liquid at 0° C., while mutton fat resulted in a depôt fat having a melting point greater than 50° C.[‡] Similarly, abnormal acids such as erucic acid have appeared in depôt fat as a result of feeding oils containing it under similar conditions.[§] Another way of showing changes of this nature is to feed animals on diets in which fat constitutes a very high proportion of the food energy. The following Table illustrates the length to which this process can be carried.[||] Rats were fed on a basal diet with the addition of various fats in amounts such that these fats constituted 60 per cent. of the total calorie intake. The iodine value of the food and depôt fats which were determined are recorded in the following Table :—

	Iodine values.	
	Food fat.	Depôt fat.
Soya bean oil	132	122
Corn oil	124	114
Cotton seed oil	108	107
Lard	63	71
Butter fat	36	55

In more physiological conditions, in which unusual fats are given in normal amounts, the depôt fat usually shows a relatively slight change in composition. Since fat is only deposited when food in excess of the energy requirements is ingested, this is to be expected, for the fat of the food will tend to be used for immediate energy purposes. Usually the lower fatty acids are not deposited in the depôt fat, but are either oxidised immediately or built up into higher fatty acids; this again illustrates the ability of the animal to maintain the depôt fat reasonably constant in composition.[¶]

It would be expected that the composition of milk fat would also be markedly affected by the addition of fat to the diet, for a cow may be producing daily several pounds of fat in its milk, while its normal food contains a relatively small amount of fat. Such is the case. In one experiment, when a cow was fed on linseed oil, the iodine number of the milk fat rose from 30, its normal figure, to 70·4. Further, after the introduction of iodised fat subcutaneously, iodised fats are found in the milk. In another experiment a bitch, which had been fed with mutton suet and had deposited in her tissues a fat of high melting point, produced a milk the iodine number of which was the same as that of the mutton suet. In this case the fat of the milk had evidently been derived from the food *viâ* the tissues, since during lactation the animal was being fed on meat which was poor in fat.

---

[*] Henriques and Hansen. *Skand. Arch. f. Physiol.*, 1901, **11**, 151.

[†] Channon and Harrison. *Biochem. J.*, 1926, **20**, 84.

[‡] Lebedeff. *Pflüger's Arch.*, 1883, **31**, 11.

[§] Munk. *Arch. f. path. Anat. u. Physiol.*, 1884, **95**, 407.

[||] Anderson and Mendel. *J. Biol. Chem.*, 1928, **76**, 729.

[¶] Eckstein. *J. Biol. Chem.*, 1929, **81**, 613; Powell. *J. Biol. Chem.*, 1930, **89**, 43.

This question of the effect of feeding with fat on the composition of depôt and milk fats is one of great importance in stock raising, since food concentrates, often the bye-products of other industries, are much used for fattening purposes. The question of flavour has also to be considered, because particular food fats may readily cause the appearance in the fat, or in the milk, of any flavour present in them.

(*b*) *The Formation of Fat from Carbohydrate.* It has long been known to farmers that animals may be fattened on a diet rich in carbohydrate and relatively poor in fat. The classical experiment demonstrating the conversion of carbohydrate into fat in the animal body was that of Lawes and Gilbert.* In this experiment one of two pigs taken from the same litter was killed and completely analysed. The other pig received a diet for a given period and the total intake of fat, carbohydrate and protein in the diet was measured. At the end of the experiment this animal also was killed and analysed and a balance sheet drawn up, in which the essential points are as follows :—

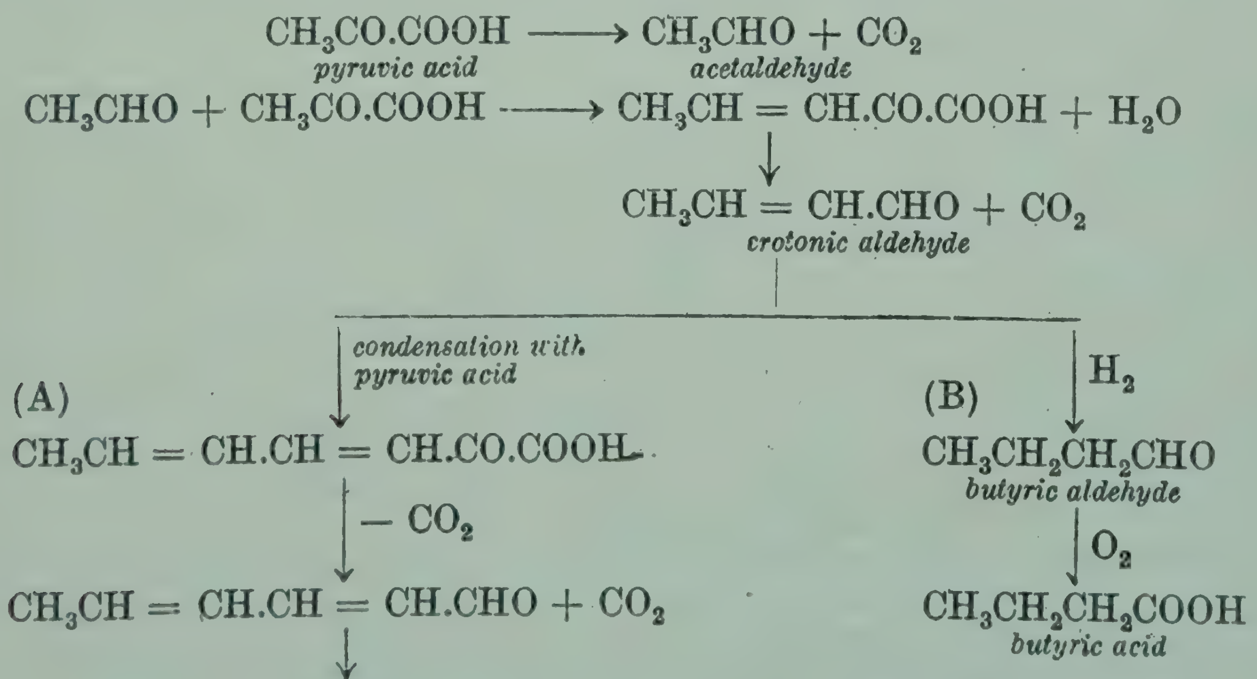
Total body fat increase .	71·2 lbs.	Protein in food . . .	64 lbs.
Fat in food	12·4 ,,	Body protein increase . .	6·5 ,,
Fat synthesised . . .	58·8 ,,	Protein available for fat synthesis	57·5 ,,
Carbon in fat synthesised .	45·3 ,,	Carbon in 57·5 lb. protein .	27·4 ,,

From these figures it is seen that whereas the synthesised fat contained 45·3 lb. of carbon, that available from protein was only 27·4 lb., *i.e.* 17·9 lb. of carbon had appeared as fat and must have been derived from carbohydrate.

Fattening of animals on a high carbohydrate diet causes the production of a harder fat than is normal,† with a corresponding decrease in its iodine value, which illustrates again the influence of food on depôt fat composition.

The means whereby sugar is converted into fatty acids is not known, although it is almost certain that it is not converted to a fatty acid by direct condensation ; an initial breakdown and a resynthesis is the probable path. Since carbohydrate metabolism both in the animal and in alcoholic fermentation passes through the three carbon stage of pyruvic acid, it is very probable that this compound forms the starting point from which fat is synthesised. It is significant that phosphate and vitamin B_1, the pyruvate co-carboxylase, are essential for conversion of carbohydrate to fat, as well as for decarboxylation of pyruvic acid to form acetaldehyde. Aldehydes possess the property of condensing together to form aldols, which may lose water with the production of double bonds. Pyruvic acid can also condense, either with itself or with an aldehyde.

Various schemes to explain the formation of fatty acid from carbohydrate have been postulated, based on this property. One such is briefly outlined below : ‡


$CH_3CO.COOH \longrightarrow CH_3CHO + CO_2$
pyruvic acid
acetaldehyde
$CH_3CHO + CH_3CO.COOH \longrightarrow CH_3CH = CH.CO.COOH + H_2O$
$CH_3CH = CH.CHO + CO_2$
crotonic aldehyde
condensation with pyruvic acid
(A)
$CH_3CH = CH.CH = CH.CO.COOH$
$- CO_2$
$CH_3CH = CH.CH = CH.CHO + CO_2$
H_2
(B)
$CH_3CH_2CH_2CHO$
butyric aldehyde
O_2
$CH_3CH_2CH_2COOH$
butyric acid

---

* *Phil. Trans. Roy. Soc.*, 1859 ; *J. Anat. and Phys.*, 1877, **11**, 577.

† Ellis and Zeller. *J. Biol. Chem.*, 1930, **8**, 185.

‡ Lubrzyrska and Smedley. *Biochem. J.*, 1913, **7**, 375.

In the above scheme a molecule of acetaldehyde derived from pyruvic acid condenses with a further molecule of pyruvic acid with the production of an unsaturated keto acid, which loses carbon dioxide to yield the unsaturated crotonic aldehyde. This latter by reduction would yield butyric aldehyde, which on oxidation would yield butyric acid (B). (It should be noted that no unsaturated aldehyde intermediaries have yet been detected.) Instead of following this path, crotonic aldehyde could condense with a further molecule of pyruvic acid and by the same procedure an unsaturated aldehyde with six carbon atoms is produced (A). This compound could either be reduced and oxidised to produce hexoic acid, or condensed further. By similar processes fatty acids containing any *even number of carbon atoms* could be synthesised and they would be saturated or unsaturated dependent on the extent to which the double bonds took up hydrogen. The fatty acids so produced would then be coupled with glycerol formed in the body from glucose, by way of glyceric aldehyde.* Such a theory would satisfy the fact that a whole series of even numbered saturated fatty acids may be produced from carbohydrate, together with unsaturated acids. Oleic acid is quantitatively by far the most important unsaturated acid synthesised from sugar, the amounts of other unsaturated acids being very small.

Adipose tissue contains a variable amount of glycogen, and it has been suggested that this represents a stage in the local conversion of carbohydrate into fat.†

(*c*) *The Formation of Fat from Protein.* The proteins are absorbed as amino-acids and we know that some 60 per cent. of the amino-acids in the food can be converted to carbohydrate in the body. It might reasonably be thought, therefore, that some of this might be converted to fat. There are, however, two limitations which make it certain that fat does not normally arise from protein. The first of these is that amino-acids can only be converted to glucose after their deamination in the liver. This is a relatively slow process, so that there is never made available at any given moment a large amount of sugar derived from amino-acids ; hence, after a high protein meal there is no pronounced rise in blood sugar. Secondly, it is difficult to satisfy the normal energy requirements of the body by ingestion of protein only. Hence any carbohydrate produced from protein would be at once oxidised for immediate energy purposes. In particular conditions, nevertheless, fat may be formed from protein, as was shown by the experiment of Atkinson, Rapport and Lusk,‡ in which large amounts of protein were given to a well-nourished animal resting in a warm environment. Some lower forms of animal life normally form fat on an exclusively protein diet, as was shown by Weinland for certain fly larvæ.

Utilisation of Depôt Fat. The power of the organism to store up reserve carbohydrate is strictly limited and probably 400 grammes represents the extreme limit of storage of carbohydrates in the body. On the other hand, in most animals there is practically no limit to the amount of fat which can be laid down by overfeeding, since the amount of reserve fat deposited will clearly depend on excess of income over expenditure. As soon as the energy expenditure exceeds the income, this fat store is drawn upon, so that adipose tissue is the one which presents the greatest loss during starvation, during which as much as 97 per cent. of it may disappear.

(*a*) *Functions of Depôt Fat.* As a reserve material, fat possesses very great advantages over carbohydrate or protein. In the first place its calorific value, 9·3, is more than twice as great as that of these other foodstuffs, 4·1. Secondly, the storage of either protein or carbohydrate entails the presence of much water, while adipose tissue is virtually pure fat (95 per cent.). Thus an enormous energy reserve may be laid down in a relatively small space. This fat in the depôts is not a static dump of material, but is constantly being turned over, even when the total amount is unchanged, or increasing. This can be shown by feeding with " labelled " fats, *e.g.* with fat containing deuterium.

Besides this function as reserve material, fat serves as a protection against temperature change and against mechanical injury, a simple illustration of the former being the apparent immunity of women, as compared with men, to

---

[*] Embden, Schmidtz and Boldes. *Biochem. Z.*, 1912, **45**, 174.

[†] Tuerkischer and Wertheimer. *J. Physiol.*, 1942, **100**, 385 ; Wertheimer, *J. Physiol.*, 1945, **103**, 359.

[‡] *J. Biol. Chem.*, 1922, **53**, 155.

exposure to cold surroundings; an immunity which women in a normal state of nutrition owe to the fact that nature has been more liberal, as art is more tolerant, of subcutaneous fat in the female than in the male.

(*b*) *The General Utilisation of Fats in the Body.* After a few days' starvation the sole sources of the energy which is available must be fats and proteins, and it is possible by an estimation of the nitrogen output to determine what fraction of the total energy evolved is to be ascribed to protein metabolism. Thus in the case of Cetti, the professional faster, it was found that the nitrogenous metabolism per unit of body weight between the fifth and tenth days of starvation corresponded to an average of 1 gramme of protein per kilogramme body weight daily. In order to oxidise this amount of protein about 1 c.c. of oxygen per minute would be required. Cetti's total oxygen consumption was 5 c.c. per kilogramme per minute, so that four-fifths of the oxygen absorbed was required for the oxidation of non-nitrogenous substances and these could only have been fats. In animals with a large store of fat the proportion of the energy obtained at the cost of the fats may be still greater.

How the fat leaves the depôts is uncertain, but there is some evidence that it leaves, as it enters, in the form of fine particles of neutral fat in the blood. During fasting these are called "*lipomicrons*"; they show periodic bursts of increase at five-hourly intervals during starvation.

The oxidation of fats supplies energy, not only for the production of heat, but also for the performance of mechanical work. Fat is found as a normal constituent of all muscle fibres and the amount of this substance is greater in proportion to the activity of the muscles concerned. Thus the heart muscle and the red muscle of the diaphragm contain larger amounts of fat than the pale voluntary muscles which have to undertake only short periods of activity.* In the human heart muscle 15 per cent. of the solids are soluble in ether and more than one-half of the ether extract is composed of fat; this would be sufficient to supply the energy of the contracting heart for six or seven hours' work. Evidence from various sources suggests that the heart does use fat to provide energy. In most instances, however, it seems that the tissues use carbohydrate in preference in order to supply energy.

In prolonged moderate exercise the respiratory quotient is the same as during rest, provided that the diet remains the same. The respiratory quotient under these conditions depends merely on the nature of the food oxidised, and it makes no difference what intermediate changes take place, such as that of fat to carbohydrates or *vice versâ*, so long as all the food used finally undergoes complete oxidation. This is well shown in the following Table, in which the oxygen consumption and respiratory quotient are compared in a man resting and working on three different diets—one principally fat, one principally carbohydrate, and the other principally protein.

Diet principally	Resting		Working		kg. m. of work done	Per kg. m. of work	
	c.c. oxygen used per min.	Resp. quotient	c.c. oxygen used per min.	Resp. quotient		c.c. oxygen used	Cal.
Fat . . .	319	0·72	1029	0·72	354	2·01	9·39
Carbohydrate .	277	0·90	1029	0·90	346	2·17	10·41
Protein . .	306	0·80	1127	0·80	345	2·38	11·35

During short bouts of severe exercise, as we have seen in Chapter X., the respiratory quotient of the *excess* metabolism is about unity, which might be interpreted as meaning that the extra energy is derived by the oxidation of carbohydrate. This is the case even when the subject is on a fat diet. But in that case after the exercise, or if hard exercise is prolonged, the quotient falls and this is interpreted as indicating that the muscle (and liver) are replenishing their stores of glycogen at the expense of fat.

---

* Leathes. *J. Physiol.*, 1904, **31**, Proc. ii.

(*c*) *The Possible Conversion of Fat into Sugar.* When a starved rat is fed on a fatty acid containing an *odd* number of carbon atoms (*e.g.* propionic acid), an accumulation of glycogen is demonstrable in the liver. But natural fats only contain fatty acids of even carbon numbers and do not produce this effect. Although the evidence is conflicting, it has been claimed that constant production of glucose is normally proceeding in the liver, even during fasting, and this glucose might arise from fat.[*] Further, the perfusion of livers with blood containing butyric acid leads to the production, not only of ketone bodies, but also, it is claimed, of glucose.[†] Also, the evidence that the energy of muscular contraction is derived largely from carbohydrate requires, as a logical necessity, the possibility, at all events during fasting, of the conversion in the body of fat into carbohydrate.

During hibernation, as Pembrey showed, the marmot has a very low respiratory quotient, which may be as small as 0·3 or 0·4. The excess intake of oxygen can be explained only by assuming that there is going on in the body a conversion such as that of fats into carbohydrates. The same conversion, accompanied by a very low respiratory quotient, has been alleged to take place in certain cases of diabetes. In many cases, when the diabetic animal is fasting, or is receiving a purely protein diet, a uniform ratio has been found to exist between the glucose and the nitrogen excreted; this ratio, $\frac{G}{N}$ generally equals 2·8. This represents a conversion into sugar of 45 per cent. of the protein used. In a few cases, however, even during complete starvation, the G : N ratio has been stated to amount to as much as 10 or 12. These animals are stated to be practically free from carbohydrates, so that the sugar excreted in the urine can come only from the breakdown of fats, since the protein could not provide all of it.

Various suggestions have been made as to the path by which fat could be converted to carbohydrate.

The Oxidation of Fatty Acids

The normal animal combusts fat to carbon dioxide and water and generally, therefore, no oxidation products are excreted in the urine. Any theory of fatty acid oxidation must account for the fact that fatty acids with any number of carbon atoms up to 20, both saturated and unsaturated, have to be oxidised. Secondly, it should be remembered that all the natural fatty acids contain an even number of carbon atoms. Possible hypotheses would be (i) direct and complete oxidation in the tissues generally, or (ii) stagewise oxidation begun in the liver and finished in tissues generally. For various reasons (ii) is most favoured. It may occur in various ways.

(*a*) *β-oxidation.* Evidence as to the nature of the intermediate steps has been obtained by administering to animals the phenyl derivatives of fatty acids and examining the urine for the end products. The tissues cannot readily disrupt the benzene ring, which must therefore be excreted, and which may carry with it a partially oxidised residue of the fatty acid molecule to which it was attached. Knoop [‡] found that when a series of phenyl fatty acids, beginning with phenyl acetic acid, were given to animals, one of two products only appeared in the urine, either phenyl formic acid (benzoic acid), or phenyl acetic acid, in accordance with the following scheme :—

---

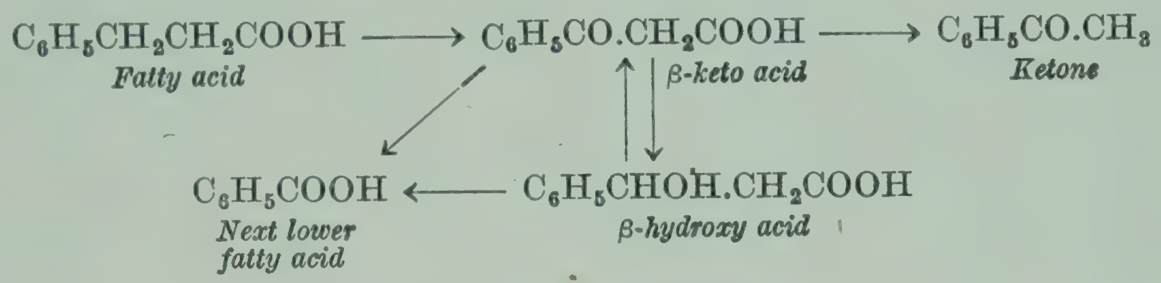
[*] Young. *J. Physiol.*, 1936, **87**, 11 P.; Soskin *et al.* *Amer. J. Physiol.*, 1935, **114**, 110.

[†] Blixenkrone-Møller. *Z. f. Physiol. Chem.*, 1938, **252**, 137.

[‡] Knoop. *Beit. Z. Chem. Physiol. Path.*, 1905, **6**, 150.

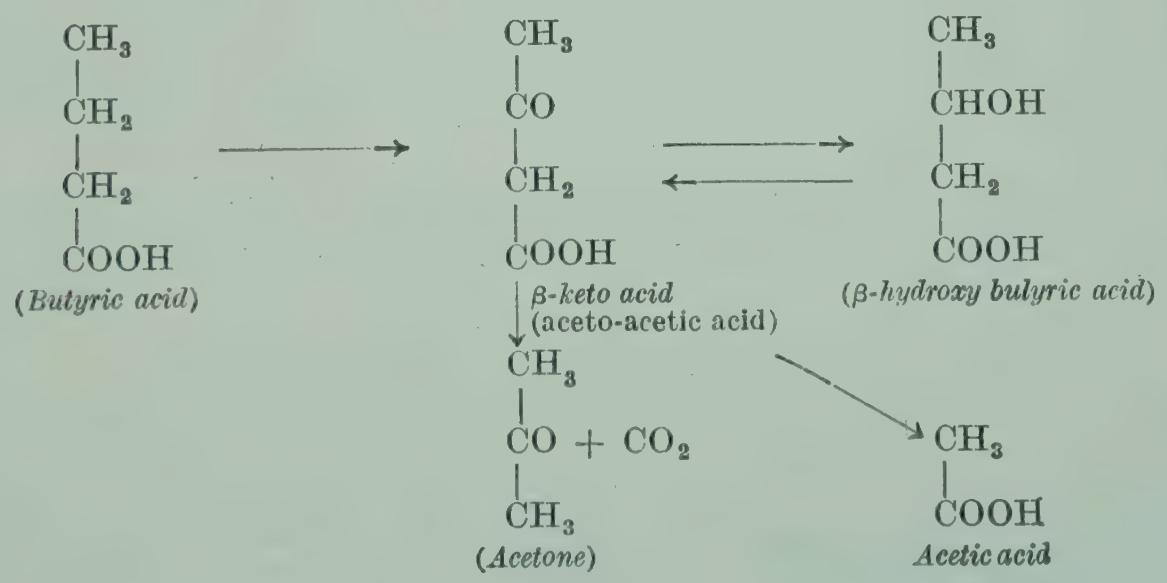
Acid administered.		No. of C atoms in fatty acid chain.	Product excreted.	No. of C atoms lost.
Phenyl formic acid (benzoic acid)	C_6H_5COOH	1	C_6H_5COOH	0
,, acetic acid	$C_6H_5CH_2COOH$	2	$C_6H_5CH_2COOH$	0
,, propionic acid	$C_6H_5(CH_2)_2COOH$	3	C_6H_5COOH	2
,, butyric acid	$C_6H_5(CH_2)_3COOH$	4	$C_6H_5CH_2COOH$	2
,, valeric acid	$C_6H_5(CH_2)_4COOH$	5	C_6H_5COOH	4

Since the number of carbon atoms lost was always an even number, this suggested that the fatty acids lost their carbon atoms two at a time, the odd-numbered acids giving rise to benzoic acid and the even-numbered to phenyl acetic acid. Dakin * showed that further intermediate products could be obtained in accordance with the following scheme :


$C_6H_5CH_2CH_2COOH$
Fatty acid
$C_6H_5CO.CH_2COOH$
β-keto acid
$C_6H_5CO.CH_3$
Ketone
C_6H_5COOH
Next lower fatty acid
$C_6H_5CHOH.CH_2COOH$
β-hydroxy acid

These two lines of evidence are the basis on which the theory of successive β-oxidation of fatty acids has been evolved. When a fatty acid is oxidised, the first product is a β keto acid, which is then oxidised at the carbonyl group to the fatty acid with two carbon atoms less. The two-carbon fragment split off may be presumed to form acetic acid. Thus, in the normal oxidation of a saturated acid like stearic acid, $C_{18}H_{36}O_2$, after the corresponding intermediate products had been first produced, palmitic acid, $C_{16}H_{32}O_2$, would be formed by loss of two carbon atoms. This process of losing two carbon atoms at a time is repeated with the production next of myristic acid, $C_{14}H_{28}O_2$, and similarly through successive stages of C_{12}, C_{10}, C_8, and C_6 to butyric acid, $C_4H_8O_2$. The higher intermediate products have not been isolated, probably because the further oxidation is rapidly and completely achieved, so that these products never accumulate.

The butyric acid, by further β-oxidation would yield aceto-acetic acid, and this then can undergo one of three further changes, reduction to β-hydroxy butyric acid, decarboxylation to acetone, or conversion to acetic acid :—


CH_3 CH_2 CH_2 COOH
(Butyric acid)
CH_3 CO CH_2 COOH
β-keto acid (aceto-acetic acid)
CH_3 CHOH CH_2 COOH
(β-hydroxy butyric acid)
CH_3 CO + CO_2 CH_3
(Acetone)
CH_3 COOH
Acetic acid

The oxidations in question appear to be carried out by the liver, and liver perfusion investigations bear this out. If fatty acids are added to blood used for perfusing a liver, there is an increased amount of acetone bodies produced

---

* Dakin. *J. Biol. Chem.*, 1908, **4,** 419 ; **5,** 173, 303.

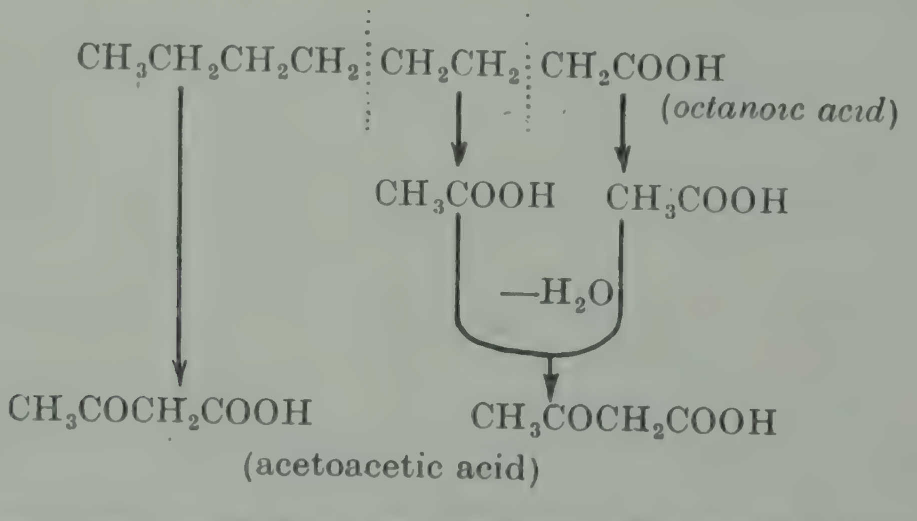
when even-numbered fatty acids are used, but not in the case of odd-numbere ones.* If aceto-acetic acid is added, β-hydroxybutyric acid is formed.

Further confirmatory evidence has come from studies of the oxidation o fatty acids added to liver slices. When lower fatty acids were added to live slices, only those which possessed an even number of carbon atoms gave ris to additional aceto-acetic acid and, further, in the liver, aceto-acetic acid wa the final oxidation product.†

There is much evidence in support of the view that β-oxidation is one of th normal mechanisms by which fatty acids are oxidised. But for the oxidatio to be normal and complete, there must be concurrent oxidation of carbohydrate In any state in which carbohydrate cannot be adequately used (diabete mellitus), or in which there is insufficiency of carbohydrate for oxidatio (starvation), excess of fat has to be burnt in order to supply sufficient energy. In this condition a state of *ketosis* occurs in which there is excretion in the urin of aceto-acetic and β-hydroxy butyric acids and acetone. These are precisely the products which, as we have seen, should be obtained on the β-oxidation theory, if, as in these conditions, fatty acid oxidation were incomplete.

The exact origin of the ketone bodies is important. According to the original β-oxidation theory the two-carbon fragments (? acetic acid) split off at each stage are oxidised to CO_2 and water, and aceto-acetic acid only arises when, at the four-carbon stage, butyric acid is in turn subjected to β-oxidation. Therefore, each molecule of fatty acid should yield only one molecule of aceto-acetic acid.

The amount of ketone bodies produced is, however, often greater than this, *e.g.* the C_8 octanoic acid can all be converted into aceto-acetic acid (two molecules). There is evidence that this is effected by β-oxidation followed by condensation of the two-carbon fragments to form aceto-acetic acid, thus :—



Proof of this condensation of 2-carbon fragments is that if octanoic acid containin C^{13} in the COOH group was incubated with liver slices, the aceto-acetic acid produced contained labelled carbon equally in the CO and COOH groups ; also, acetic acid with C^{13} in the COOH incubated with kidney tissue produced the same doubly labelled aceto-acetic acid.‡

(*b*) *Multiple Alternate Oxidation.* Another way in which ketone bodies could arise has been proposed. In successive β-oxidation, the end-products would be, from each molecule of fatty acid, one of ketonic acid, and several of acetic acid, *e.g.* from palmitic acid :—

$$\text{Palmitic acid} + 6{\cdot}5\ O_2 = 1\ \text{ketone} + 6\ \text{acetic acid.}$$

But there is the possibility that the fatty acid molecule may be simultaneously attacked at different points, with the production only of several molecules of ketone bodies, *e.g.* :—

$$\text{Palmitic acid} + 1{\cdot}5\ O_2 = 4\ \text{ketone.}$$

By determining the oxygen consumed, and the ketone bodies produced per molecule of fatty acid destroyed, it should be possible to decide between these possibilities. This was done by Stadie, and it was found that 1·1 ± 0·2 moles oxygen were used, and that about 4 moles of ketone resulted, when the experiment was carried out on liver slices. Moreover, no acetic acid, or any C_2 to C_8 acid could be found : the low R.Q. (0·3) showed that these had not been first formed and then oxidised.

---

* EMBDEN, SALOMON and SCHMIDT. *Beit. Z. Chem. Physiol. Path.*, 1906, **8**, 129 ; EMBDEN and MARX, *ibid.*, 1908, **11**, 318.

† QUASTEL and WHEATLEY. *Biochem. J.*, 1933, **27**, 1753 ; JOWETT and QUASTEL, *ibid.*, 1935, **29**, 2143, 2159, 2181.

‡ WEINHOUSE *et al.* *J. Biol. Chem.*, 1944, **155**, 143 ; 1945, **157**, 751.

(c) ω-oxidation.* It was observed that on administration of the triglycerides of some fatty acids, dicarboxylic acids with the same number of carbon atoms could be isolated from the urine, *i.e.* the terminal methyl group of the fatty acid had been oxidised to a carboxyl group. The biggest yield of such products was obtained with the triglyceride of capric acid, $CH_3(CH_2)_8COOH$, which gave rise in the urine to sebacic acid, $HOOC(CH_2)_8COOH$, with smaller quantities of the corresponding dicarboxyl acids with six and eight carbon atoms, *i.e.* suberic acid, $HOOC(CH_2)_6COOH$, and adipic acid, $HOOC(CH_2)_4COOH$.

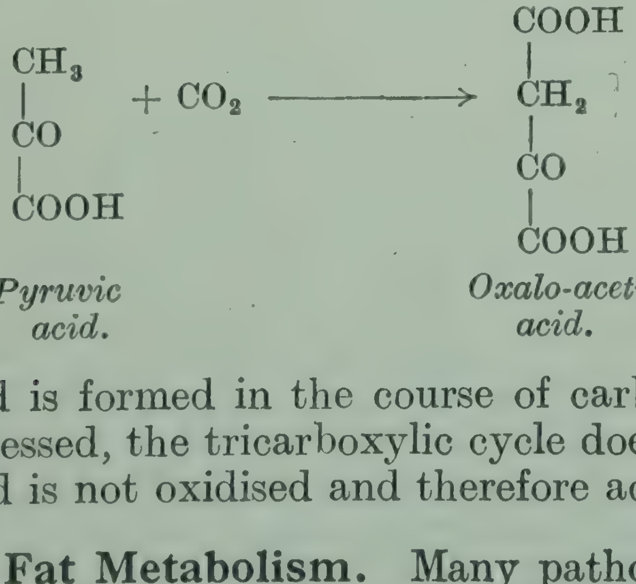
The fate of the ketone bodies which are formed in the course of fat metabolism varies. Some tissues, *e.g.* muscle and kidney, can oxidise them at reasonable rates, but if large amounts are produced their disposal becomes difficult, especially if carbohydrate is lacking (a condition which entails an accelerated production of ketones).

The Anti-ketogenic Action of Carbohydrate. The ketone bodies do not appear in the blood and urine when adequate carbohydrate or other glucose-yielding (anti-ketogenic) substance is being burnt at the same time, and this is the basis of the expression that fat is burnt in a " fire of carbohydrate."

This anti-ketogenic action of carbohydrate has been investigated by dietary studies in which mild ketonuria has been induced by a restricted diet. Such studies † show that when the molecular ratio of ketogenic to anti-ketogenic substances in the food exceeds two, acetone bodies appear in the urine. The anti-ketogenic substances will, of course, include not only dietary carbohydrate, but sugar which may be derived from glycerol, or from the straight-chain amino acids such as alanine ; whilst the ketogenic fraction will include, not only the fatty acids, but also those amino-acids, such as tyrosine, which give rise in their metabolism to aceto-acetic acid. These results are clearly of importance from the point of view of diabetic feeding, and simply expressed they mean that fat oxidation begins to be incomplete when the amount of " fat " in the diet is more than twice that of the carbohydrate actual and potential.

Ketosis is produced or augmented by the injections of certain extracts of the anterior pituitary body. Similar claims regarding the suprarenal cortex have also been made (Ch. L.).

There have been various suggestions for explaining why concurrent oxidation of carbohydrate is essential for preventing ketosis. An attractive one can be understood by reference to the tricarboxylic acid cycle (p. 143). As a result of its operation, acetic acid can be oxidised to CO_2 and water, instead of being condensed to form aceto-acetic acid, and any aceto-acetic acid formed in other ways can also be oxidised. But an essential step in the cycle is oxalo-acetic acid, and this is produced from pyruvic acid and CO_2 under the influence of β-carboxylase :—



$$\begin{array}{l} CH_3 \\ | \\ CO \\ | \\ COOH \end{array} + CO_2 \longrightarrow \begin{array}{l} COOH \\ | \\ CH_2 \\ | \\ CO \\ | \\ COOH \end{array}$$

Pyruvic acid. *Oxalo-acetic acid.*

The pyruvic acid is formed in the course of carbohydrate oxidation, and hence, if this is depressed, the tricarboxylic cycle does not function efficiently, and aceto-acetic acid is not oxidised and therefore accumulates.

The Liver and Fat Metabolism. Many pathological and experimental findings indicate that the liver plays a special part in fat metabolism. The most striking fact in this connection is that fat is transported to the liver in many conditions. Normally the liver contains some 3 per cent. of lecithin and 1 per cent. of glyceride. In the variety of conditions which may be classed together as fatty livers, large amounts of glyceride may appear, increases up to 70 times the normal amount having been observed. These fatty livers appear in the early stages of starvation, in poisoning by phlorrhizin, chloroform, alcohol, phosphorus and a variety of other chemical agents ; they occur after pancreatectomy, in clinical and experimental diabetes, in pernicious anæmia and a number of other pathological states. They may be induced also by diets high in fat or containing cholesterol. Although transportation of fat to such tissues as the heart and kidney also occurs, its frequency is very much less, as is also the actual degree of fat infiltration.

---

* Verkade and Lee. *Biochem. J.*, 1934, **28**, 31.

† Schaffer. *J. Biol. Chem.*, 1921, **47**, 433, 449 ; 1921, **49**, 143 ; 1922, **54**, 399.

Investigations to determine the source of this fat, by experiments in which animals have been starved to remove the depôt fat, then fattened on a "labelled" fat, *e.g.* on fat containing deuterium instead of some of its hydrogen atoms,* and finally poisoned with carbon tetrachloride, or starved, to cause deposition of fat in the liver, indicate that the mobilised liver fat has come from the depôts. But if "labelled" fat containing deuterium is given until the depôts contain that fat, and the animal is then given a diet low in protein but rich in carbohydrate, the liver fat contains no deuterium and so has probably been formed from carbohydrate. Hence it appears that new-formed fat appears directly in the liver, but mobilised fat is transported thither, so that the problem resolves itself into a consideration of why depôt fat appears in the liver. It is to be recalled that the fatty acids present in the liver fats have a higher iodine value than have those of depôt fat, but not much more than half that of the acids present in the liver as phosphatide. This and other findings led Leathes † to suggest that depôt fatty acids are mobilised and transferred to the liver, where they undergo a process of desaturation and conversion to phosphatides. They are then carried to the tissues for oxidation in this now more unstable form or else are, at least partially, oxidised in the liver itself. The proofs that ketone bodies are formed in the liver, including the fact that after removal of the liver the ketosis following pancreatectomy or injection of the pituitary ketogenic fraction is reduced or abolished, supports the view that the initial stages of the oxidation of fats takes place in the liver. There is a close relationship between the rate at which fat is transported to the liver and the extent of ketosis. Thus, adrenalectomy reduces both ketosis and the fat of the liver induced by injection of anterior pituitary extracts. If we accept the view that the oxidation of fats proceeds in the liver up to the stage at which 4-carbon ketone bodies are produced, then it would seem probable that these substances, set free into the blood stream, are finally oxidised, together with glucose—also released by the liver—in the tissues.‡ That certain tissues, *e.g.* heart muscle, can utilise β-hydroxybutyric acid has been demonstrated beyond question.§

Although the evidence in favour of this particular type of desaturation has lost much of its force, this must not be regarded as disposing of desaturation as a biological process, and recent work seems to suggest that desaturation does in fact occur.‖ Incubation and other experiments with livers, both normal and fatty, suggest that the liver can desaturate fatty acids, and it is tentatively suggested that this desaturation, which may be carried out by other tissues also occurs in the α–β positions. We must therefore conclude that desaturation probably occurs, although the original evidence and the particular type o desaturation which was originally put forward is largely disproved.

The common C_{18} acids in nature are oleic, linoleic and linolenic acids. These products presumably built up in the plant from carbohydrate, have their double bonds in the 9 : 10

---

* Barrett, Best and Ridout. *J. Physiol.*, 1938, **93**, 367.
† Leathes and Raper. "The Fats," London, 1925.
‡ Blixenkrone-Møller. *Z. f. physiol. chem.*, 1938, **252**, 117.
§ Waters *et al.* *Amer. J. Physiol.*, 1938, **122**, 542.
‖ For references see Artom, "Annual Review of Biochemistry," 1934–35, 212.

9 : 10. 12 : 13 ; 9 : 10. 12 : 13. 15 : 16 ; positions respectively. The occurrence of these double bonds symmetrically disposed, three carbon atoms from each other, might suggest that they have been evolved by a synthesis from a three-carbon unit, which has been produced from sugar by some such means as has been indicated in an earlier paragraph. Animals on a fat-free diet build up much oleic acid with traces of higher unsaturated acids (arachidonic acid), although the ability to synthesise linoleic acid is absent, while linolenic acid does not occur. Linoleic acid is a dietary essential for normal growth.*

After prolonged fat-starvation, the per cent. of fat in the dry tissues is as high as in normal animals, showing that formation of saturated fats and olein have proceeded as usual. It remains for future research to show why oleic acid with one double bond can be apparently synthesised with ease, whereas linoleic acid with two double bonds cannot.

Lipotropic Action. Investigation of the fatty liver problem received considerable impetus by the discovery by Best and his colleagues of the control of liver fat—the so-called *lipotropic action*, exerted by choline,† which is now recognised as an essential dietary constituent. On diets which are free from choline, large amounts of fat appear in the liver, whereas if choline be present in the diet, as normally it always is, there is no accumulation of liver fat. Furthermore, in experiments in which fatty livers have been produced it is possible to cause the disappearance of the fat by the administration of choline, whereas in the absence of choline the liver fat remains at its abnormally high level. The kidneys also undergo fatty changes, with extensive destruction of tubules and glomeruli. This action of choline in liver fat control is clearly of great importance, because it enables us by dietary means to produce fatty livers ; furthermore, it allows the amount of liver fat to be controlled at will by the inclusion of suitable amounts of choline in the diet ; it has also provided a new method whereby the *rôle* of the liver in fat metabolism may be further investigated. The amounts of choline which are effective may be quite small, for in the rat 5 milligrammes per day has been found to prevent a rise of some 12 per cent. of fat, expressed as the wet liver weight.

The mechanism of this choline action is probably that it provides the base choline for phosphatide synthesis, and so links up with the hypothesis of lecithin as a carrier of fatty acids.

Proteins also exert a lipotropic action,‡ which appears to be due to their content of methionine.

The effect of choline in preventing the fatty liver was discovered through the observation that depancreatised dogs developed fatty livers. Extracts of pancreas contain a specific substance, *lipocaic*. This may be an enzyme liberating methionine (from protein), which then is used for trans-methylative formation (p. 29) of choline.§

Fatty Degeneration. In certain pathological conditions various organs of the body such as the liver, undergo what has been described as *fatty degeneration* and this was interpreted as a conversion of the proteins of the cells into fat. We have seen already that such a conversion does not occur in the normal animal, and so-called fatty degeneration is due to a mixture of two contributing factors. The first of these is *fatty infiltration, i.e.,* mobilisation from the depôts occurs, as has already been discussed. Secondly, the protoplasm of cells contains " fat " in such a condition that its presence cannot be detected by histological means, probably because it is in some form of adsorption complex with the proteins. Hence the presence or absence of visible fatty globules affords very little clue to the quantities of fat in the cells. In various diseases and toxic conditions this fat is unmasked, due to the effect of various agents in breaking down the intimate cell structure and this is the second factor involved in fatty degeneration. The expression " *unmasked fat* " used in histology and pathology refers to such protoplasmic fat.

The Rôle of Phospholipides in the Tissues. On the grounds of its chemical structure (70 per cent fatty acid) and of the peculiar physical property of forming almost true solutions in water, lecithin obviously suggests itself as one of the means whereby fatty acids could be transported in the body. This brings us to consider whether phospholipides are integral parts of cell structure, metabolically inert from the point of view of fatty acid metabolism, or whether they are actively concerned in the latter process, or both.

---

* Burr *et al.* *J. Biol. Chem.*, 1932, **97**, 1.

† Best. *Lancet*, 1934, 1274 ; Best and Huntsman. *J. Physiol.*, **83**, 255.

‡ Channon *et al.* *Bioch. J.*, 1938, **32**, 976 ; 1943, **37**, 483.

§ Chaikoff *et al.* *J. Biol. Chem.* 1945, **160**, 489.

There is considerable evidence that formation of phospholipides may occur in the intestine, from fatty acids; when labelled fatty acids were given they were traceable in the phospholipides of the intestinal mucosa after a short interval,* and it has been suggested that fatty acids may thus be transported in part into the portal blood and not, like neutral fat, into the lymph.

The total ether-soluble material present in such tissues as the normal liver, heart, kidney, etc., consists largely, not of glycerides, but of a mixture of phosphatides. If these two groups of constituents be separated from one another and the fatty acids prepared from them by hydrolysis, the phosphatide acids are always found to be more unsaturated than the glyceride acids, and these in turn are usually more unsaturated than those of the depôt fat. Some illustrative figures are as follows:

	Iodine values of:		
	Total fatty acids.	**Fatty acids of phosphatide.**	**Fatty acids of glyceride.**
Liver	113	134	76
Heart	107	128	67
Kidney	92	117	84
Spleen	90	98	75
Lung	66	77	60

The total amount of ether-soluble material in such tissues as those given in the table varies between some 2 and 4 per cent. and, depending on dietary conditions to some extent, one-half to two-thirds or more may consist of phosphatide with the remainder as glyceride, cholesterol and cholesteryl esters.

The problem then is to decide why phospholipides occur in all tissues and why they are more unsaturated than are the glycerides. Much work has been done to attempt to determine whether lecithin is actively involved in fat metabolism as such. As we have seen when dealing with absorption, there is *primâ facie* evidence that formation of lecithin in the intestine, at any rate, may play an active part either in the absorption of fat or in the processes immediately following it. Further, in certain conditions, the amount of liver phosphatide appears to be changed after a fatty meal, and its fatty acids modified by the fat of the diet; thus iodised fats appear both in the liver and blood corpuscle phosphatides after administration.† Further, the effect of choline ‡ in controlling the level of liver fat also suggests an active *rôle* for lecithin.

On the other hand, the phosphatide of most tissues shows very little change in extreme dietary variations, and such tissue as brain, which contains much phosphatide, apparently does not utilise fats. Moreover, it is particularly the unsaturated fatty acids which are built up into phosphatides in the intestine, and which are retained with tenacity in the tissues. We may conclude then by saying that, although the chemical and physical properties of lecithin make it suitable from the point of view of fat transport and oxidation, the clear evidence so far adduced in favour of its active participation is relatively slight, save perhaps in the case of the liver and intestinal mucosa.

This brings us then to the consideration of the structural functions of phospholipides. These are essential cell constituents, forming an integral part of the cell membrane, and it might be anticipated, therefore, that even though these substances contain 70 per cent. of a fatty acid in their molecule and therefore have a high calorific value, they could not be called upon as reserve material without the destruction of the cells in which they are present. That this is so is indicated by much work carried out by Maeyer and Terroine.* The French School of workers make use of the expression the "*constant element*," by which they mean the fatty acids present in one form or another which cannot be called upon in time of need. That this constant element exists has been shown by studying the lipoids of animals which have died from inanition ; the larger part of the fatty acids of these are present as phosphatides and are constant both in amount and composition. This would imply that the phosphatide fatty acids of a given tissue are characteristic of that tissue and cannot be varied by diet. Whilst the constancy of amount seems to have been confirmed by further work, it has been shown that the nature of the fatty acids, particularly of intestinal mucosa and liver, may be substantially modified by diet. It is, however, interesting to note that the degree of unsaturation of phosphatide fatty acids does seem to be more or less characteristic of the tissue in that the iodine value of the acids in the different tissues seems to run in the same order.

---

* SINCLAIR. *J. Biol. Chem.*, 1936, **115**, 211.

† ARTOM. *Arch. inter. Physiol.*, 1933, **36**, 101; ARTOM and PERETTI. *Arch. inter. Physiol.*, 1933, **36**, 351.

‡ BEST and HUNTSMAN. *J. Physiol.*, 1935, **83**, 255.

As to the part taken by this essential fat of the working cells, in the life of the cell, our knowledge is still very scanty. Fatty acids enter into the constitution of the phospholipides and glycolipides which form important constituents of the protoplasm, Golgi apparatus and limiting membrane of every cell. As constituents of the membrane itself, fatty substances therefore have a protective action and also regulate the passage of substances into the cell across the membrane. The presence of the phospholipides as integral constituents of all protoplasm suggests that they may play a part in the normal transformations which occur within the cell, and may represent, so to speak, the currency into which fat is transformed in order to participate in the vital processes, and to be transported and utilised for the needs of the cell. Much work also indicates that there is a physiological antagonism between cholesterol and lecithin, which presumably is related to their physical properties.

Changes during Absorption. As we have seen earlier, the fatty acids absorbed in the small intestine are partly resynthesised in their passage through the intestinal epithelium, some to neutral fat and some to phospholipide. The fat resulting from this resynthesis may, for several reasons, be of a different character from that of the food. In the first place, the ease of absorption of fatty acids varies with their physical properties. Unsaturated liquid acids are very readily absorbed, especially to form phospholipide, whilst saturated acids are less well absorbed with increasing length of the carbon chain, so that the higher saturated acids are poorly absorbed unless unsaturated acids are being absorbed at the same time. Hence, in some conditions, higher fatty acids may not be absorbed, but passed into the fæces. Further, bile and other secretions contain small amounts of fatty acids and these too may be absorbed with the food fat and cause some modification in its condition.

The original food fatty acids appear in the chyle, and eventually in the blood, as a mixture of phosphatide, glyceride and cholesteryl esters, glyceride usually being in a vastly greater proportion.

This brings us to consider whether the various lipides of the blood show any changes after a meal. Here the evidence is not conclusive. There is, of course, an increase in the neutral fat in the blood, and lipæmia may develop. With this there is also an increase in phosphatide and cholesteryl ester at different times after absorption, and the extent of these changes is different in corpuscles and plasma. Attempts have been made to interpret the findings as meaning that the phosphatides or cholesteryl esters play an active part in the transport of fatty acids † ; another possible view is that with the arrival of food fatty acids into the blood, considerations of physico-chemical equilibrium may require readjustment of the proportions of the phosphatide and cholesteryl ester present. Since cholesterol has a melting point of 150° C., while its unsaturated esters are liquid at body temperature, it may be that these unsaturated esters occur on this account ; but at this stage the somewhat conflicting evidence would make conclusions premature.

---

* TERROINE. "Physiologie des substances grasses." Paris, 1919.

† BLOOR. *J. Biol. Chem.*, 1915, **23,** 317 ; 1916, **24,** 447.

Post-absorptive blood fat in the case of some animals consists of phosphatide and cholesteryl esters, with very little glyceride, and in this case it is found that the fatty acids of the cholesteryl esters have a higher iodine value than do those of the phosphatide.* The significance of this has yet to be established.

Sterol Metabolism. The metabolic fate of cholesterol and the other sterols is poorly understood at the present time. Cholesterol can be synthesised in the body † and is a constituent of all cells. It is absorbed from the intestine, provided there is sufficient fat in the diet, and the absorption depends also, of course, on the presence of bile. Plant and some other sterols are absorbed poorly, if at all, from the alimentary tract. Bile always contains cholesterol, and the excretion of sterol in the fæces is largely in the form of coprosterol. Just as fatty acids are converted to neutral fat during absorption, so cholesterol is also in part esterified with fatty acids. The ability of the animal to destroy cholesterol appears to be limited, because if 1 per cent. of cholesterol is present in the diet of rats or rabbits, large quantities are deposited in the different organs, mainly in ester form, and intense fatty livers in which enormous quantities of glyceride also occur may be produced by such means. Further investigation of this problem may have clinical significance. Deposition of cholesterol in the tissues seems to be general, although brain is an exception. In such experiments as these part of the cholesterol is, however, oxidised to unknown compounds. Various bile acids which are being continuously formed are almost certainly derived from cholesterol, as are the sex hormones.

For Reference

ANDERSON and WILLIAMS (1937). The Rôle of Fat in the Diet. *Physiol. Rev.*, **17**, 335.
BLOCH (1947). The Metabolism of Acetic Acid in Animal Tissues. *Physiol. Rev.*, **27**, 574.
BLOOR (1939). Fat Transport in the Animal Body. *Physiol. Rev.*, **19**, 557.
BLOOR (1943). Biochemistry of the Fatty Acids. (New York : Reinhold.)
BROBECK (1946). Mechanism of the Development of Obesity in Animals with Hypothalamic Lesions. *Physiol. Rev.*, **26**, 541.
BURR and BARNES (1943). Non-caloric Functions of Dietary Fats. *Physiol. Rev.*, **23**, 256.
COWGILL (1945). Relative Nutritive Values of Animal and Vegetable Fats. *Physiol. Rev.*, **25**, 664.
FRAZER (1940). Fat Absorption and its Relationship to Fat Metabolism. *Physiol. Rev.*, **20**, 561.
FRAZER (1946). The Absorption of Triglyceride Fat from the Intestine. *Physiol. Rev.*, **26**, 103.
MACLEAN (1943). The Metabolism of Fat. (London : Methuen.)
MCHENRY and PATTERSON (1944). Lipotropic Factors. *Physiol. Rev.*, **24**, 128.
NEWBURGH and CONN (1944). Obesity. *Physiol. Rev.*, **24**, 18, 31.
SINCLAIR (1934). Physiology of the Phospholipids. *Physiol. Rev.*, **14**, 351.
STADIE (1945). The Intermediary Metabolism of Fatty Acids. *Physiol. Rev.*, **25**, 395.
THANNHAUSER and SCHMIDT (1946). Lipins and Lipidoses. *Physiol. Rev.*, **26**, 275.
WOOD (1946). The Fixation of Carbon Dioxide and the Inter-relationships of the Tricarboxylic Acid Cycle. *Physiol. Rev.*, **26**, 198.
WERTHEIMER and SHAPIRO (1948). The Physiology of Adipose Tissue. *Physiol. Rev.*, **28**, 451.

4. CARBOHYDRATE METABOLISM

The Blood Sugar

The carbohydrates present in the body consist of the sugar in the blood and tissue fluids, and the glycogen stored in the tissues. All the carbohydrates in an ordinary diet are ultimately transformed in the alimentary tract into glucose, fructose and galactose. These three, together with mannose, are utilisable from the blood, and conversion into glucose takes place in the body. Glucose is certainly present in the blood ‡ (60–150 mg./100 ml. in man), but whether it is the only simple carbohydrate there is not definitely settled ; no others have so far been isolated from ordinary blood. In the fœtus, however, fructose may be present and account for one-third of the sugar (sheep).

Accurate determination of sugar in the blood is difficult ; blood contains reducing substances other than sugars which may be equivalent to no less than 27 mg./100 ml. of glucose,* but when the same method is used the error due to "non-sugar" substances is constant for a given species, and when it is deducted from the total "apparent glucose," the "true glucose" is obtained. In clinical work, results are often given in terms of apparent glucose, and are 4–30 mg. too high, according to the method of estimation used. The non-sugar reducing substances (mainly glutathione and thioneine) are principally contained in the corpuscles, 100 ml. of which contain 47 mg. (expressed as glucose), while plasma only contains 10 mg./100 ml.† In consequence of this, the corpuscles in normal blood have often been stated to have a glucose concentration 1·1 times that of serum, whereas in fact the ratio is only 0·77.

---

* BLOOR. *J. Biol. Chem.*, 1924, **59**, 543 ; CHANNON and COLLINSON. *Biochem. J.*, 1929, **23**, 1212.

† CHANNON. *Biochem. J.*, 1925, **19**, 424.

‡ WINTER. *Biochem. J.*, 1930, **24**, 851.

The glucose of the blood plasma is in equilibrium with that in the tissue fluids. Although the body only has, say, 2·5 l. of blood plasma, which would contain about 2·5 g. of glucose, the tissue fluid amounts to about 15 l., and so represents a considerable store of glucose—at least 15 g. Consequently if glucose is suddenly added to the circulating blood, as by intravenous injection, the blood sugar is not raised so much as might be expected, since, after a few minutes, glucose has diffused out into the tissue fluid until equilibrium is reached. If, for instance, 50 g. of glucose were given intravenously, the blood sugar would not rise to 1 per cent., but only to about 0·35 per cent. Conversely the withdrawal by the tissues of 2 g. of glucose would not cause the blood sugar to disappear, but would scarcely lower it appreciably.

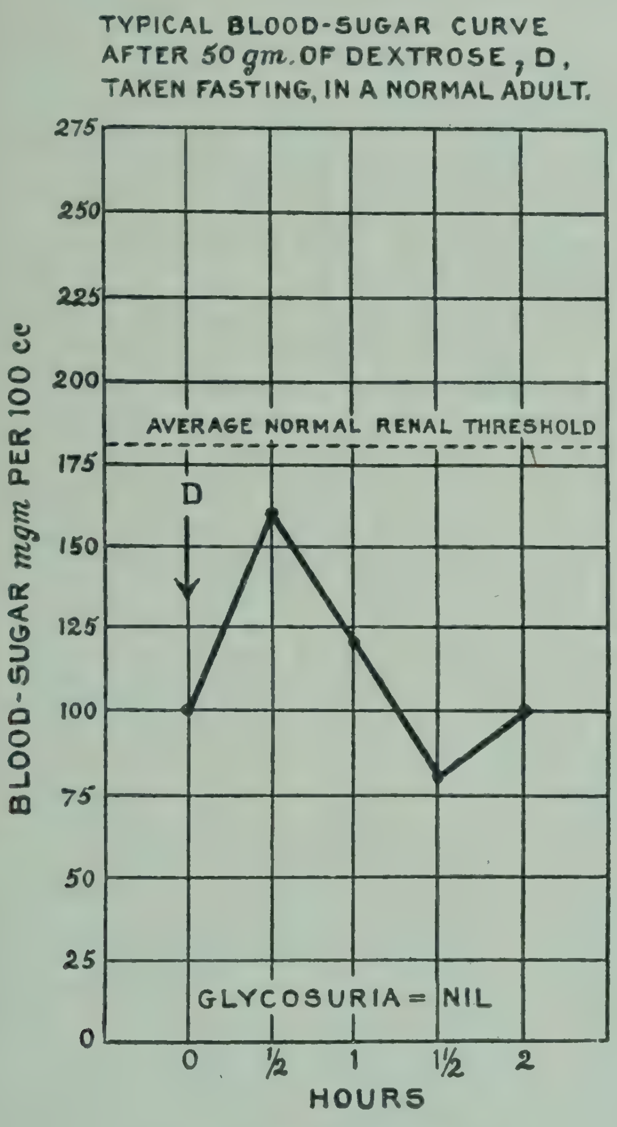


Fig. 589. (Harrison.)

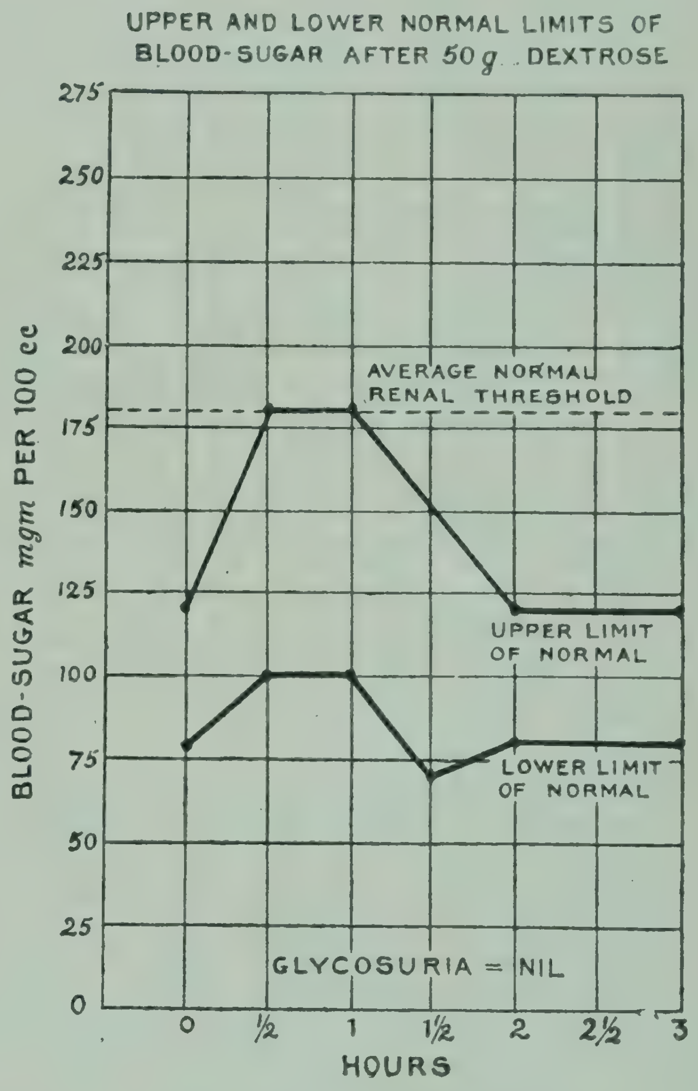


Fig. 590. (Harrison.)

The sugar-values in Figs. 589 and 590 are for total reducing substances, as usually determined for clinical purposes. True glucose figures are 20–30 mg. lower.

During the absorption of a large carbohydrate meal, considerable quantities of monosaccharides enter the blood, and the amount of sugar in the blood is increased. When the absorption is rapid, as, for example, after giving 50 g. of glucose to a fasting person, the systemic blood sugar is, in fact, raised for a while to about double the fasting level; the portal blood sugar is higher still, and runs closely parallel with the systemic blood sugar. This procedure of giving glucose and estimating the blood sugar at half-hourly intervals for the next two or three hours is called the *glucose tolerance test.**

---

* Somogyi. *J. Biol. Chem.*, 1927, **75**, 33; Herbert and Bourne. *B.M.J.*, 1931, i., 94.

† Somogyi. *J. Biol. Chem.*, 1928, **78**, 117.

For this purpose the fasting level of the blood sugar is determined ; it is usually from 60 to 90 mg./100 ml. (true sugar). Then 50 g. of glucose dissolved in water is taken by the mouth. Samples of blood are drawn each half hour for three hours, and their sugar content determined. Within thirty minutes after taking the glucose the true sugar will be found to have risen to a level not usually exceeding 150 mg./100 ml. (Figs. 589, 590) ; in one hour the amount has begun to fall, and at the end of one and a half to two hours the blood sugar has returned to normal again, or it may be even a little below (Fig. 589). The " blood sugar curves " so obtained do not vary greatly in normal persons, and are of considerable value clinically in investigating disorders of carbohydrate metabolism, such as diabetes. It is usual to test the urine for glucose at intervals after taking the glucose ; none should be present. If an ordinary meal containing insoluble carbohydrates capable of digestion, such as starch, be taken instead of glucose, the rise is of slightly longer duration, owing to the relatively slow formation and absorption of the products of digestion (Fig. 591).

It is important to note that although such alterations of the blood sugar occur, these are normally never very great. There is a normal range for each species. In ruminant animals the level is low and the range narrow, *e.g.* for sheep, 30–60 mg./100 ml. ; for most other mammals about 70–150 mg./100 ml. This is because there is a well-controlled means for regulation of the blood sugar concentration, *i.e.* of the sugar available to the tissues. Even during prolonged starvation glucose does not disappear from the blood, nor even fall notably below the amount present after a fast of only a few hours. Conversely, even after the heaviest carbohydrate meal, the systemic true blood sugar in normal subjects does not rise above about 150 mg./100 ml.

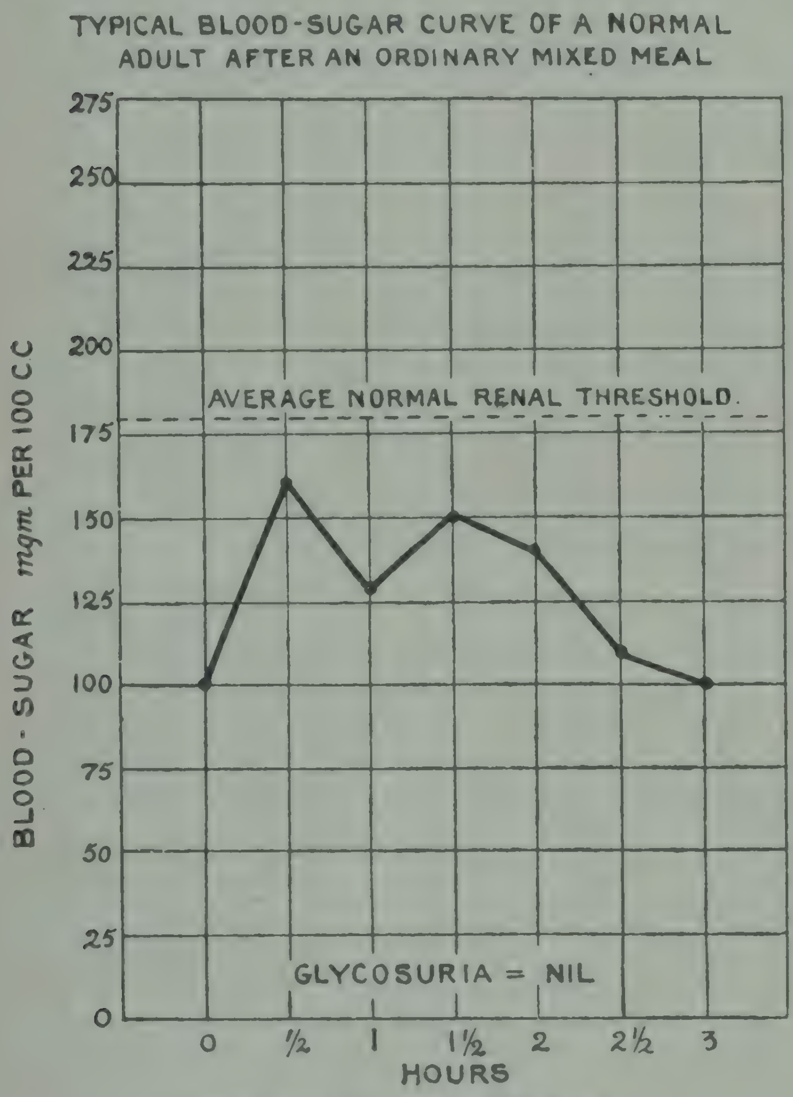


FIG. 591. (HARRISON.)

This relatively steady content of the blood sugar depends on an adjustment of the rates at which sugar enters and leaves the blood stream. If the blood sugar remains constant, it is because, like the water level in a tank, the rates of inflow and outflow are equal. If one is greater than the other, the level must rise or fall *continually* until they again become equal. Under normal conditions the regulating mechanism is " set " to a true glucose level of about 70 mg./100 ml., but under abnormal conditions it may be set at higher or lower levels ; when this is the case we cannot infer that the rates of entry or exit of glucose have undergone any prolonged change—all we can say is that the regulation is set at a higher or lower level.

The following scheme summarises briefly the sources of gain and loss of sugar to the blood :—

---

* *v.* HARRISON. "Chemical Methods in Clinical Medicine." (Churchill.) 1947, p. 133.

Gain.	*Loss.*
From absorbed sugar.	Removed by diffusion into tissue fluid.
From liver glycogen.	Destroyed in blood (glycolysis).
Added by diffusion from tissue fluid.	Removed by liver and other tissues.
	Removed by excretion.

The blood sugar can arise directly from two sources only, viz., from absorbed carbohydrate or from liver glycogen. In the fasting animal the only direct source of blood sugar is the glycogen of the liver.

One method by which glucose may be removed from the blood and tissues is by *glycolysis.* This is due to the action of enzymes by which the glucose is converted into lactic acid. The enzymes are present in the leucocytes, and also in various tissues. The red cells contain diphosphoglycerates and adenosine triphosphate, both of which can yield phosphate ions.

A first stage in glycolysis is brought about by the enzyme *hexokinase,* which catalyses the reaction :—

Glucose + adenosine triphosphate ⟶ glucose (6) phosphate
+ adenosine diphosphate

the subsequent stages being as in the formation of lactic acid from hexose phosphate derived from glycogen (*v.* p. 142). Glycolysis occurs in the circulating blood, particularly in its transit through the lungs,* and accounts for the constant presence of lactic acid in the blood. The glycolysis in shed blood may be prevented by the addition of small amounts of sodium fluoride.†

Sugar is constantly being removed from the circulating blood for use by the cells of the tissues. Hence we find that the venous blood from most parts contains less glucose than is present in the arterial blood, the difference being usually 4–5 mg./100 ml.‡ Since the sugar concentration of the arterial blood is approximately constant, it follows that, even in starvation, glucose must be regularly added to the blood at about the same rate as that at which it is removed. When glucose is being absorbed from the alimentary canal, much of it is rapidly removed from the blood, some of it being burned and some deposited in the form of glycogen in the tissues. Nevertheless, it is not usually removed quite as rapidly as it enters, so that the systemic blood sugar does show a slight rise ; this tends to bring about a still faster rate of removal, so that by the time the blood sugar content has reached 150 mg./100 ml. deposition and combustion keep pace with absorption. In this way many normal persons can dispose of 200 or 300 g. of carbohydrate at a meal. Should the rate of removal be unable, however, to keep pace with the rate of absorption into the blood, or, should glucose be *rapidly injected* into the blood stream, then the excess of glucose is removed from the blood stream by excretion by the kidneys, and appears in the urine, a condition called *glycosuria.* When sugar is excreted in the urine its concentration therein is always much greater than that in the blood ; if the sugar in the blood amounts to, say, 400 mg./100 ml., the urine may contain 5 or 6 per cent. of glucose.

Glucose taken up from the blood by the tissue cells may suffer one of three fates : combustion, conversion to glycogen, or conversion to fat. The last-named process, when carbohydrate metabolism is normal, perhaps plays a larger part in disposing of glucose than has been supposed. By the use of isotopes it has been shown § that in a particular instance, in rats, of ingested glucose, only 3 per cent. is converted to glucose, while 30 per cent. goes to production of fatty acids, most of which are then oxidised. Conversions to glycogen and oxidation are considered below.

---

* Evans, Hsu and Kosaka. *J. Physiol.,* 1934, **82,** 41.

† Lovatt Evans. *J. Physiol.,* 1922, **56,** 146.

‡ Somogyi. *J. Biol. Chem.,* 1948, **174,** 189.

§ Stetten. *J. Amer. Med. Assoc.,* 1946, **132,** 373

Although the body only stores about one day's supply of carbohydrate, lowering of the blood sugar beyond a certain level cannot be effected by complete deprivation of food, however prolonged. This means that the starving animal is able to produce blood sugar from some source within the body. The most readily available source is the glycogen store, especially that in the liver, and this is first called upon in starvation. Yet, even after prolonged starvation, the liver is found to be not quite free from glycogen, though the amount present in it may be only 0·20 per cent. or less ; the muscle glycogen likewise never disappears.

If for any exceptional reason the true sugar content of the blood be brought much below 50 mg. per 100 ml., severe symptoms result, and ultimately, if the fall continues, death occurs. This condition of *hypoglycæmia* may be brought about by the action of insulin. In the rabbit an injection of insulin sufficient to cause reduction of the blood sugar to about 5 mg./100 ml. causes severe convulsive seizures, followed by coma, fall of temperature and death from respiratory failure. In man, subjective symptoms are induced when the blood sugar falls to 55 mg./100 ml. These symptoms are hunger, a sense of fatigue and anxiety, vasomotor phenomena, such as pallor or flushing, and general sweating. As the blood sugar falls still further, mental disturbances, delirium and finally profound coma supervene. By that time the true blood sugar may be down to 5 mg./100 ml. (total reducing power = 30 mg. glucose/100 ml.). These serious consequences show that a certain minimal quantity of glucose must always be present in the blood and tissues for the maintenance of the normal functions of the central nervous system.

For Reference

DORFMAN (1943). Pathways of Glycolysis. *Physiol. Rev.*, **23**, 124.
SOSKIN (1941). The Blood Sugar. *Physiol. Rev.*, **21**, 140.

STORAGE AND WITHDRAWAL OF GLYCOGEN

Liver Glycogen. It was discovered by Claude Bernard * that the blood of the hepatic veins nearly always, but particularly in fasting animals, contains more sugar than is present in the blood taken simultaneously from the portal vein, *i.e.* the liver adds sugar to the blood which passes through it. That the liver can produce glucose he showed by a simple experiment :—

The liver removed from an animal and kept warm for some time was found to contain considerable amounts of glucose. He next performed the classical experiment with the "washed liver" † ; a liver was removed from an animal immediately after death and its vessels washed out with ice-cold water or salt solution ; the tissue was then cut up and thrown into boiling water, in which it was thoroughly ground up. The extract, after separation of the coagulable proteins, contained practically no sugar—hence glucose as such was not stored in the liver. The aqueous extract, however, contained a substance precipitable by alcohol, to which he gave the name of *glycogen*, *i.e.* sugar-former. The large amount of sugar found in the liver which had been kept warm after removal was due to conversion of the glycogen into glucose by the action of enzymes present in the liver. That it is an enzyme action is proved by the fact that fresh liver may be dehydrated with alcohol, dried and powdered, and kept for months in this condition without any alteration occurring in its glycogen. If, however, the dried powdered liver be moistened and allowed to remain at the temperature of the body for some hours, its glycogen disappears and the equivalent quantity of glucose takes its place ; this change does not occur if the liver is first boiled.

In the conversion of glycogen to glucose three enzymes are concerned (*v.* pp. 95, 97), viz. *phosphorylase*, *phosphoglucomutase* and *hexokinase*, which catalyse reactions as shown on next page :—

---

* BERNARD. *Mém. Soc. Biol.*, 1849, 121 ; *Compt. Rend.*, 1850, **21**, 571.
† BERNARD. *Compt. Rend.*, 1855, **41**, 461.

(1) $\text{glycogen} + \text{phosphate} \xrightleftharpoons{(\textit{phosphorylase})} \text{glucose-1-phosphate}$

(2) $\text{glucose-1-phosphate} \xrightleftharpoons{(\textit{phosphoglucomutase})} \text{glucose-6-phosphate}$

(3) $\text{glucose-6-phosphate} + \text{A.D.P.} \xrightleftharpoons{(\textit{hexokinase})} \text{glucose} + \text{A.T.P.}$*

By reversal of the first reaction, glycogen can be produced from glucose-1-phosphate.†

Glycogen (*v.* p. 38) is a white, tasteless and odourless polysaccharide of the empirical formula $C_6H_{10}O_5$. Like starch, it resists the action of boiling potash,‡ but is hydrolysed by acids, or by amylase, into dextrins, maltose and finally glucose. There is no evidence, however, that tissue glycogen is broken down in that way. Its aqueous solution with iodine gives a brownish-red colour, and it is precipitated by 55 per cent. alcohol, strong acetic acid, and various other precipitants which abstract water.

Glycogen may be estimated by Pflüger's method § ; the principle of the method is that weighed samples of the tissue are dropped into their own volume of 60 per cent. boiling potash solution and then heated on a water-bath for one to four hours. This dissolves practically everything in the tissue ; alcohol is then added in amount to make up to 70 per cent. ; the precipitated glycogen is dissolved in water, reprecipitated by alcohol, and finally hydrolysed for three hours in 2·2 per cent. HCl. The glucose thus formed is then estimated, and the results usually expressed in terms of glucose.

Liver Glycogen Formation. If an animal is fed for twelve to twenty-four hours previously on a diet rich in carbohydrates, glycogen may be present in amount up to 12 per cent. or more of the weight of the fresh liver. Glycogen can also be formed on a diet of proteins and the liver can continue to supply sugar to the blood on a diet which includes only proteins and fats. It is doubtful whether a normal existence is possible in man in the total absence of carbohydrates from the food, though in the Eskimos the amount of carbohydrate consumed is small in comparison with the fats and proteins. It does not seem to be possible for the liver to *store* glycogen out of fats. In experiments on feeding with fats it is found that, if fats be administered after the liver has been nearly freed from glycogen, although the liver may store up fats, it does not store up any glycogen. This may be because fat is only slowly converted, so that all the glucose formed from it would be required for use forthwith instead of being stored as glycogen.

The glycogen of the liver represents a reserve material, and is the first to be called upon to meet the increased needs of the tissues during functional activity, such as muscular work or greater heat production. If an animal be starved, the glycogen quickly diminishes in the liver, although, even at the end of ten or twelve days' complete deprivation of food, small traces of glycogen may still be found in this organ. If starvation be combined with hard work, strychnine convulsions, or exposure to cold, the liver glycogen is lowered more quickly. On account of the importance of glycogen as a reserve material, it is stored up in almost all growing tissues, *e.g.* in the placenta during a certain period, in fœtal muscles, in yeast, in oysters, and in the muscles of the body generally. In fœtal muscles it may amount to as much as 40 per cent. of the total dried solids.

It is principally to Claude Bernard that we owe the conception of the *glycogenic function of the liver* which is still current. According to this, the liver cells continually form glycogen, which they can manufacture from a variety of materials. The glycogen so formed may either remain and accumulate in the liver cells, or it may be converted into glucose, and pass out into the blood. Both processes, manufacture of glycogen and conversion of glycogen to glucose, are usually in operation together; if glycogen formation predominates over breakdown, the glycogen store in the liver increases, if the formation of glucose predominates, the store is depleted and formation of glucose ultimately slows down. The balance between the formation and breakdown of glycogen in the liver is determined, on the one hand by the supply of carbohydrate available, whether derived from ingested carbohydrate or by new formation, and on the other hand by the demand for carbohydrate by the tissues of the body.

---

* CORI. *Endocrinol.*, 1940, **26,** 285.

† CORI, CORI and SCHMIDT. 1939, *J. Biol. Chem.*, **129,** 629.

‡ BERNARD. *Compt. Rend.*, 1857, **44,** 578.

§ PFLÜGER. Art. "Glycogène" in *Richet's Dictionnaire de Physiol.*, 1907, Vol. 7, [G]228.

In fasting, the blood will have a lowered content of glucose. The liver now, by the activity of its enzymes, rapidly converts its glycogen into glucose—a process called *glycogenolysis*—and this glucose is carried away in the blood of the hepatic veins which now contains much more glucose than that in the portal vein or general circulation. At the same time, new formation (*glyconeogenesis*) of glycogen from protein and fat is accelerated so as to supply the constant demand for blood sugar. In this way, the liver acts as a storehouse, in which carbohydrate, much or all of which it has made, can be temporarily stored and brought out as required.

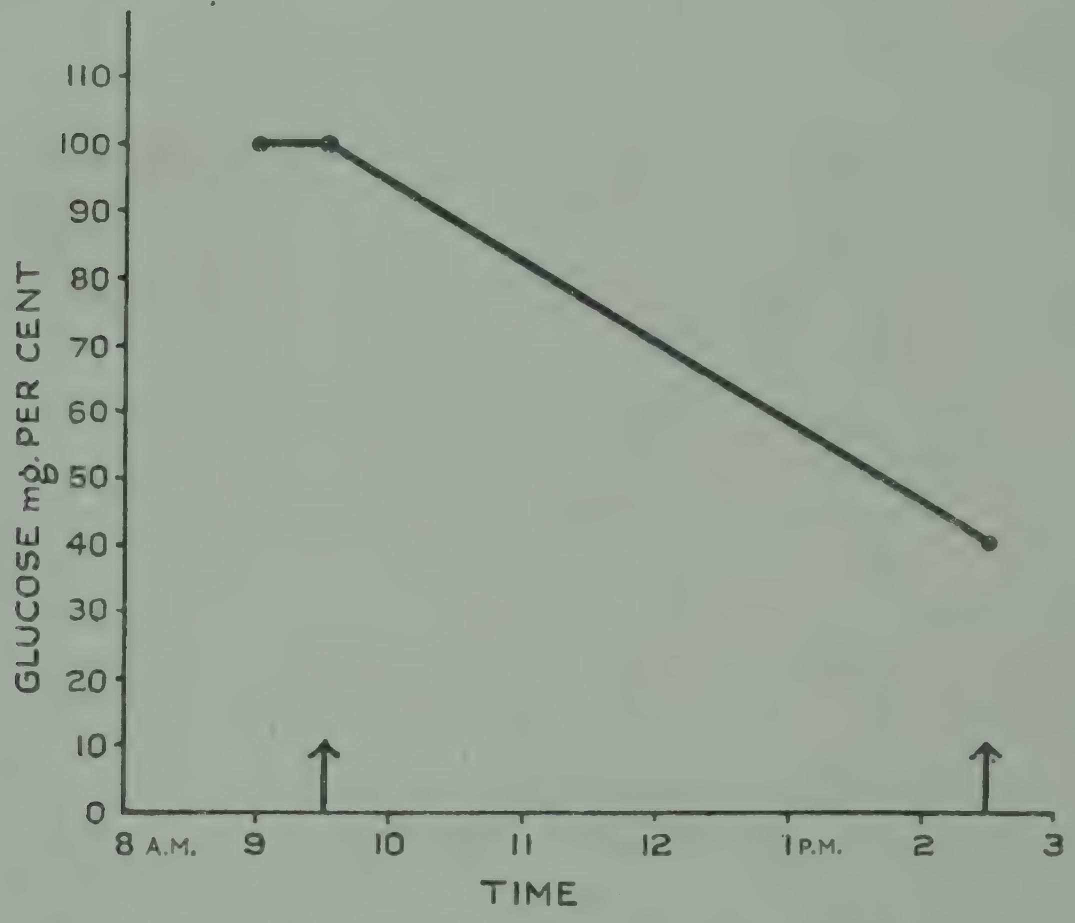


Fig. 592. Blood glucose concentrations in hepatectomised dogs (average of 12 expts.) Liver removed at first arrow; animal moribund, with hypoglycæmic sugar level at second arrow, five hours later. (MANN and MAGATH, 1923.)

The liver is giving out glucose to the blood most of the time, though not always at the same rate. In the fasting animal the liver is the sole source of blood sugar. After removal of the liver the blood sugar rapidly falls (Fig. 592), and soon the animal passes into hypoglycæmic convulsions, but this may be prevented by constant infusion of glucose at a proper rate; when the liver is present the glucose of the blood is kept up during fasting by conversion of liver glycogen to glucose at the same rate. When a carbohydrate meal is given, absorbed glucose enters the portal blood; the liver does not now add so much glucose to the blood, and if absorption is rapid enough it may add none at all. In fact, if the blood sugar is greatly raised, as by intravenous injection of glucose, the liver may actually take up sugar from the blood (Fig. 593), though this does not easily occur. The liver responds to altered blood sugar level in such a way as to tend to restore this to normal, and acts like a regulator "set" to maintain a constant blood sugar concentration: by this homoiostatic response the blood sugar is virtually self-regulating. At blood sugar levels of about 90–100 mg./100 ml., production and liberation of glucose are about equal, provided that the liver is "set," by endocrine factors, to regulate to that level. It is probable that this regulating function of the liver is dependent upon the law of mass action.

After heavy carbohydrate feeding, glycogen accumulates rapidly in the liver, and this led to the belief that the absorbed carbohydrate has been directly stored in the liver as glycogen. This is strengthened by the demonstration that, *in vitro*, glucose-1-phosphate can be converted to glycogen by phosphorylase of liver. But this simple explanation does not always suffice, because the hepatic vein blood usually contains more sugar than the portal blood, though the excess is much smaller during the absorption of a carbohydrate meal.* What happens under these circumstances is simply that more of the glycogen produced by glyconeogenesis remains in the liver, and less of it is converted into blood sugar. At higher blood sugar levels, all the glycogenetic glycogen is stored, and at still higher levels blood sugar is taken up by the liver and stored as glycogen. One substance, viz. lactic acid, is, however, regularly removed from the blood as it passes through the liver, and is converted by the liver into glycogen.

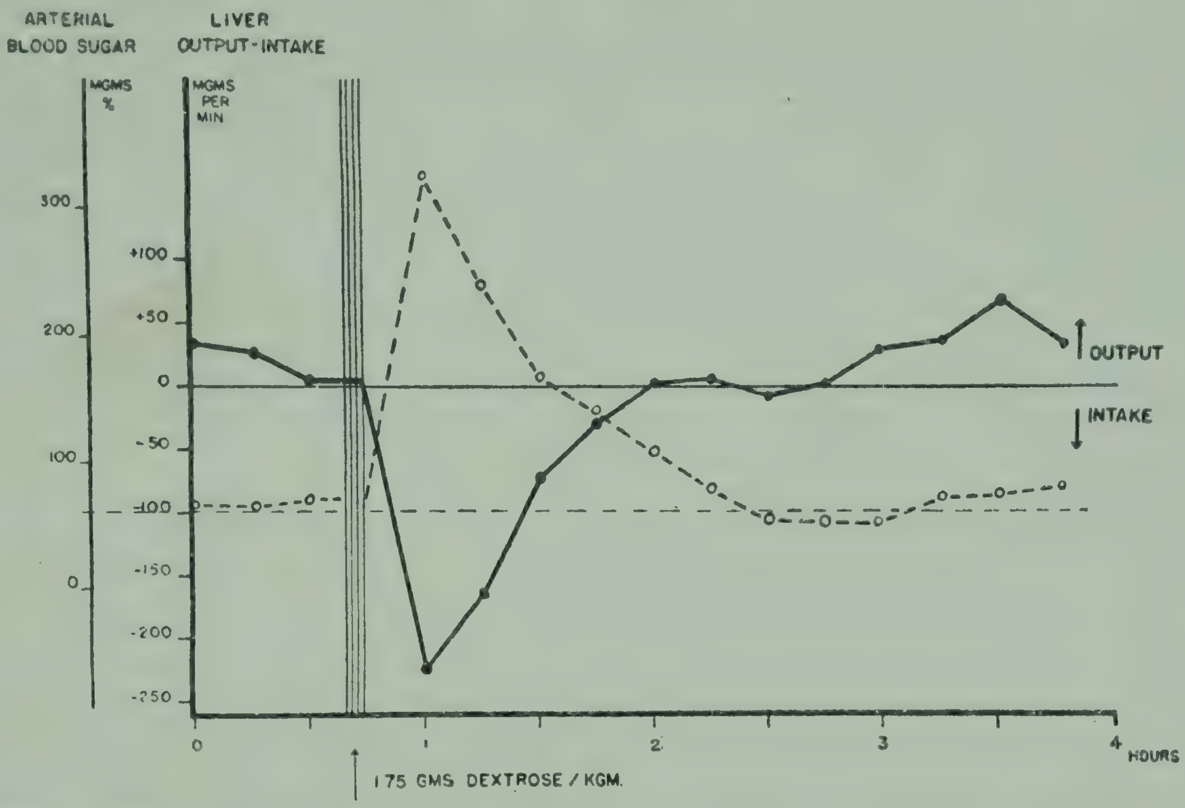


FIG. 593. Arterial blood sugar concentration (disçontinuous line), and liver output or intake of sugar (continuous line) before and after intravenous injection of glucose. Except when arterial sugar is high, the liver is giving out sugar. (SOSKIN *et al.*, 1939.)

When a liver is perfused with blood containing glucose, no glycogen is deposited, but if lactic acid, or fructose, is added instead of glucose, then this is taken up and glycogen is formed.† Probably, therefore, glucose must first be converted to something else before the liver can form glycogen from it. The liver, however, can, and does, convert pyruvic acid, glycerol, dihydroxyacetone, methyl glyoxal, and some amino-acids into glycogen. Hence the liver glycogen may not only be formed after carbohydrate ingestion, but it may also arise from protein and perhaps from fat, by synthetic processes which occur in the liver.

The glycogen content of the liver, therefore, represents the balance between glycogen formation from these various sources, and glycogen breakdown to glucose, and both these processes are continually in operation. The rate at which glycogen is formed in the liver depends mainly on the concentration of sugar in the blood ; the higher the blood sugar the less need there is for the liver to supply it, and the greater the amount of glycogen retained in the liver. When glucose is given and the liver glycogen is increased, it does not necessarily follow that this represents storage as glycogen of the absorbed glucose. Some of the accumulated liver glycogen arises by processes of glyconeogenesis and is retained by the liver, while the absorbed glucose is used by the tissues, instead of these drawing upon liver glycogen.

---

* TSAI and YI. *Chinese J. Physiol.*, 1936, **10**, 105.

† LUNDSGAARD, NIELSEN and ORSKOV. *Skand. arch. Physiol.*, 1936, **13**, 296.

There is some evidence, from experiments with C. isotopes, that carbon dioxide can be incorporated in the glycogen molecule in the liver, but not in muscle, when glucose is converted to glycogen ; such a synthesis would involve a partial oxidation of the glucose, and subsequent assembly of fragments, together with CO_2, to build up glycogen*.

There are various ways by which the process of glycogen storage in the liver may be studied. One of the simplest and most direct of these is the method introduced by Cori and Cori.† They showed that the livers of different white rats of a standard, inbred strain, contain equal amounts of glycogen when the animals are all treated identically. If the animals fast for twenty-four hours the liver glycogen is reduced to about 0·2 to 0·5 per cent. When a sugar solution, *e.g.* glucose, is given in sufficient amount, it is quickly absorbed, and at a constant rate which, as we have already learnt, is characteristic of the particular sugar. Thus, three batches of fasted rats may be taken ; to the first is given by stomach tube some water, to the second the same amount of water containing glucose, and to the third water and a similar weight of fructose. Then animals in each batch are killed each hour and the glycogen in the livers determined. It was found that when 1 to 2 g. of sugar was given per 100 g. of body weight, the maximum amount of liver glycogen was formed after four hours. After glucose the liver contained 5·3 per cent., and after fructose 5·7 per cent. glycogen. Different sugars, however, are absorbed at different rates, *e.g.* in the four hours 1·06 g. of the glucose and 0·540 g. of the fructose was absorbed per 100 g. of rat, and when the glycogen formation is recalculated, taking this into account, it was found that in four hours the amounts of glycogen deposited represented 18 per cent. of the absorbed glucose and 38 per cent. of the absorbed fructose.

Glycogen Storage in Muscle. Skeletal muscle contains from 0·05 to 1·0 per cent. of glycogen, which is formed from the blood sugar, *viâ* glucose-6-phosphate and glucose-1-phosphate by reversal of the reactions given on p. 957. The deposition of glycogen is probably always going on to replace the glycogen continually being used up by the muscle. Hence the venous blood going away from a muscle contains less glucose than the arterial blood. Though normally muscle can take up blood sugar and convert it into glycogen, even *in vitro*, this process happens much more rapidly when insulin is present in excess in the circulating blood. Although, as we have already seen, the muscles *can* transform *d*-lactic acid into glycogen, the process is a roundabout one (*v.* p. 142), and they appear able to do this much less rapidly than the liver, so that if lactic acid is set free into the circulation from the muscles during muscular exercise, it is probably mostly converted into glycogen by the liver ; the muscles then restore their glycogen from the blood sugar, which is replenished either by absorbed glucose from the alimentary canal or by breakdown of liver glycogen.

Though the muscles contain a smaller percentage of glycogen than is present in the liver, the total amount of muscle in the body is so great that the whole musculature usually stores much more glycogen than the liver. Thus, in a well-fed man of average weight the liver can perhaps hold 200 g. of glycogen, while the muscles might hold 350 g. of glycogen ; in the fasting individual the muscle glycogen would not be very greatly reduced, whereas the total liver glycogen might be as low as 10 or 20 g. It is evident, therefore, that the liver glycogen is much more mobile than the muscle glycogen, and even if the latter is considerably reduced during exercise it is soon replenished, largely at the expense of the liver glycogen.

Withdrawal of Glycogen from Liver and Muscles. It is a curious fact that glycogen is removed from liver and muscles in different forms ; it leaves the liver in the form of glucose, and the muscles in the form of lactic acid, and these may enter the blood stream. The explanation is that different enzyme systems are concerned in the two cases. In the liver, as we have seen, three enzymes between them hydrolyse the glycogen to glucose. In muscle, on the other hand, we have to deal with the complex system of enzymes capable of converting glycogen first into pyruvic acid and then into lactic acid (p. 142).

---

* Solomon *et al.* *J. Biol. Chem.*, 1941, **140,** 171.

† Cori and Cori. *J. Biol. Chem.*, 1926, **70,** 577.

The liver, then, can convert lactic acid and other substances, including glucose, into glycogen, and glycogen then into glucose ; muscle can convert glucose (or lactic acid) into glycogen, and the glycogen into lactic acid. These specialised conversions render possible an interesting cycle of changes, by means of which muscle and liver glycogen become virtually interchangeable in the body. This is seen most clearly in the action of adrenaline in certain animals, such as the rat. If rats fasted for twenty-four hours be given a subcutaneous injection of 0·2 mg. adrenaline per 100 g. of body weight, they may excrete much more glucose than could have been derived from all the liver glycogen. Yet it is found that the liver glycogen, which at first fell, at the end of three hours has *risen* from an average of 6 mg. per 100 g. of body weight to 42 mg. per 100 g. body weight, but the muscle glycogen has *fallen* from 128 mg. to 71 mg. per 100 g. body weight.* What happens under these conditions is that the adrenaline for some reason causes conversion of muscle glycogen to lactic acid, the concentration of which in the blood is at first considerably increased ; the lactic acid is then taken up by the liver and converted into liver glycogen (Fig. 594). This, "the Cori cycle," is represented thus :—

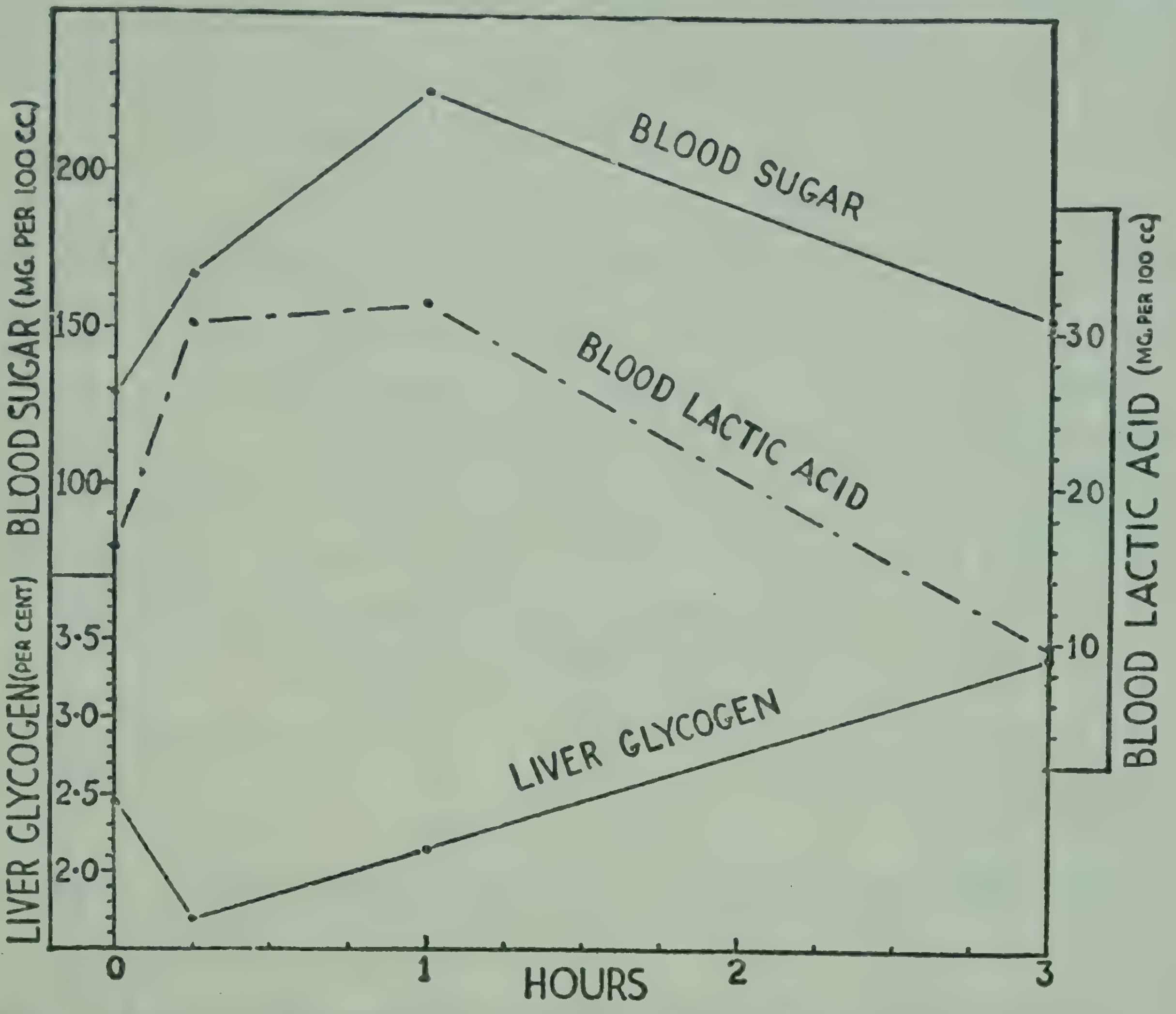
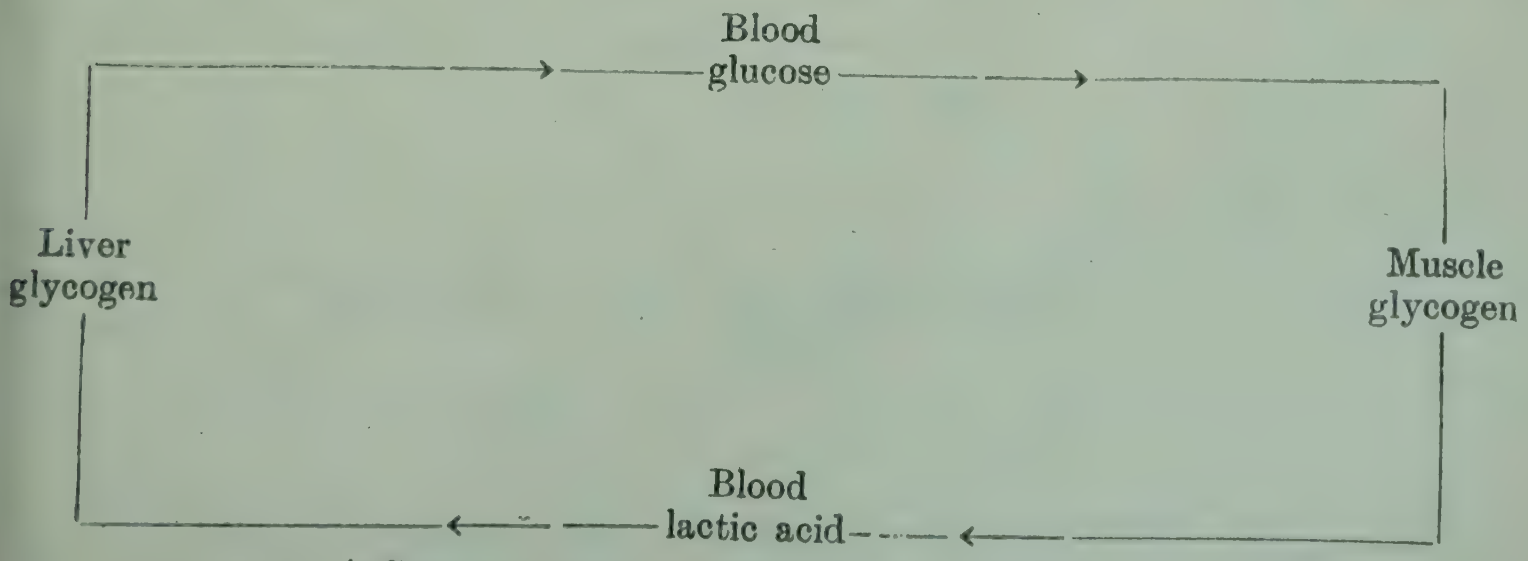


FIG. 594. Effect of subcutaneous injection of adrenaline on blood sugar, blood lactate and liver glycogen of the rat. Adrenaline 0·2 mg./100 grammes given at 0. (THORPE'S *Biochemistry*, from data by CORI and CORI, *J. Biol. Chem.*, 1930, **86**, 375.)



---

* CORI and CORI. *J. Biol. Chem.*, 1928, **79**, 309.

After sixteen hours the increased liver glycogen has disappeared again.

The principal conditions under which glycogen is withdrawn from muscle (either by complete oxidation, or as lactic acid) are under the action of severe exercise or of adrenaline.

The liver converts its glycogen into glucose under a great variety of conditions, most of which can be brought together as being due to the action upon the liver, either of nervous influences, or of adrenaline circulating in the blood. The breakdown of liver glycogen by the action of adrenaline or by nerve stimulation may be inhibited by the action of insulin. This does not imply, however, that insulin by itself might be able to lead to formation of liver glycogen.

INTERMEDIARY METABOLISM OF CARBOHYDRATE

We have already studied the inter-relations of carbohydrates, such as glucose and glycogen, in the body (pp. 38, 97), and the chemical relationships of substances closely connected with carbohydrates (pp. 31, 34). We have also considered the biochemical changes concerned in the utilisation and ultimate oxidation of carbohydrates in muscle (pp. 142, 143), and the nature of oxidising enzymes (pp. 98, 99), and tissue oxidations in general (p. 745).

It will only be necessary at this stage to bring all these data together by way of recapitulation. The following diagram does this in outline.

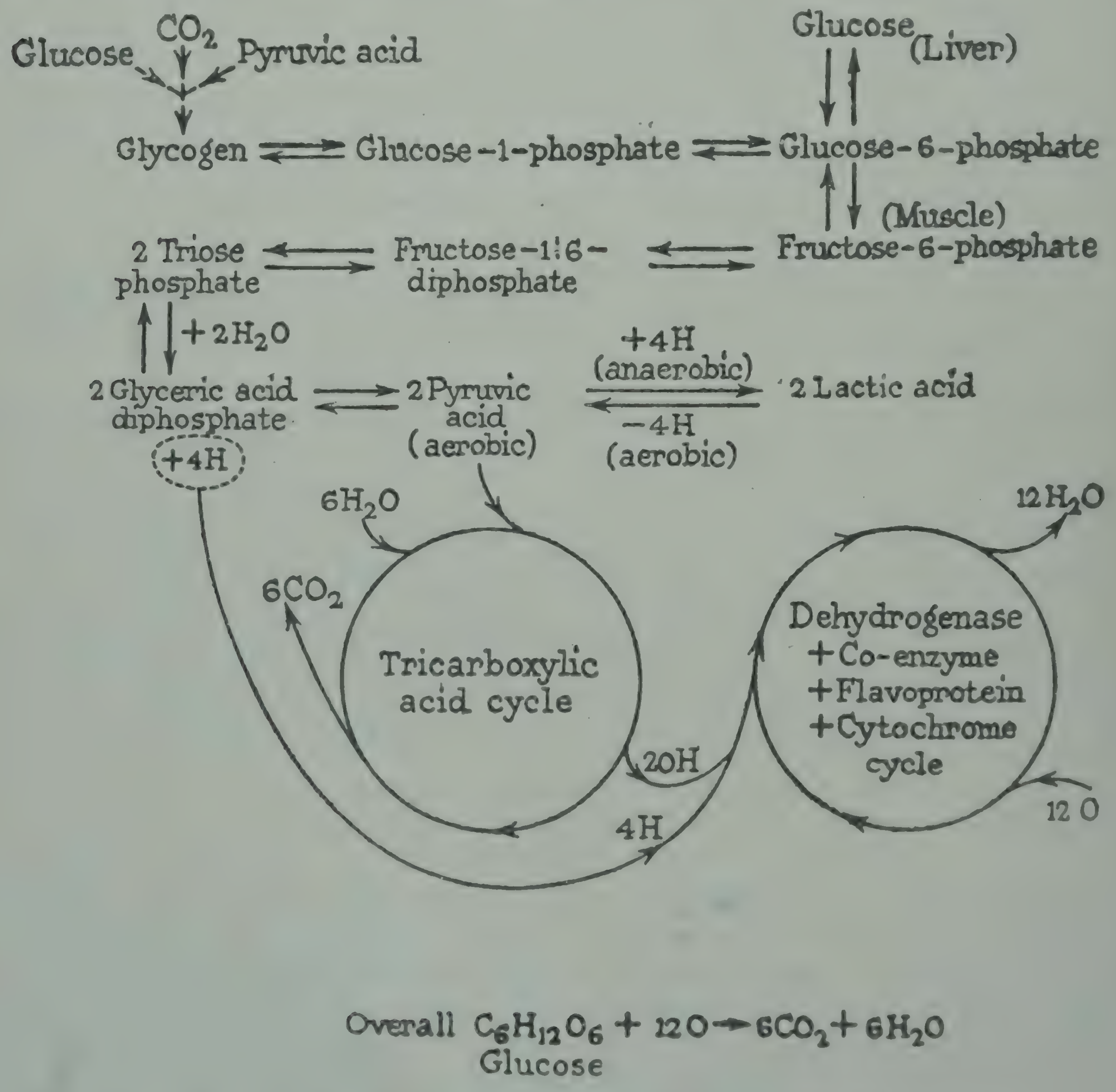

Glucose
CO_2
Pyruvic acid
Glucose
(Liver)
Glycogen
Glucose-1-phosphate
Glucose-6-phosphate
(Muscle)
2 Triose phosphate
Fructose-1:6-diphosphate
Fructose-6-phosphate
$+2H_2O$
$+4H$
(anaerobic)
2 Glyceric acid diphosphate
2 Pyruvic acid (aerobic)
2 Lactic acid
$-4H$
(aerobic)
(+4H)
$6H_2O$
$12H_2O$
$6CO_2$
Tricarboxylic acid cycle
Dehydrogenase +Co-enzyme +Flavoprotein +Cytochrome cycle
20H
4H
12 O
Overall $C_6H_{12}O_6 + 12O \rightarrow 6CO_2 + 6H_2O$
Glucose

FIG. 595. Summary of the metabolic fate of glucose and glycogen. (YOUNG. *Sci. Prog.*, 1948, **141**, 13.)

The diagram shows by what chemical stages glucose and glycogen ar interconvertible, how glucose can be converted into pyruvic acid, and how th latter substance is either oxidised to CO_2 and water in the tricarboxylic cycle (p. 143) or alternatively reduced to lactic acid. The cytochrome cycle on the right is here represented as a means by which hydrogen derived from dehydrogenase actions is ultimately oxidised by oxygen, as described on p. 747.

The importance of pyruvic acid as a substance in common in the intermediary metabolism of carbohydrate, protein and fat has already been mentioned. It should also be noted that it can be converted into either oxaloacetic acid or α-ketoglutaric acid, which enter into the tricarboxylic cycle.

HYPERGLYCÆMIA AND GLYCOSURIA

Normal urine contains a small proportion of reducing substance, but at least half of this is not glucose. The term *glycosuria* is not employed unless glucose appears in quantities large enough to give a reaction with Fehling's solution. Such a condition may easily be brought about by the injection of glucose, subcutaneously or intravenously. This causes a high blood sugar, or *hyperglycæmia*. Cane sugar or lactose, introduced in the circulation, is quantitatively excreted in the urine. Glucose, fructose, mannose or galactose, if introduced slowly enough into the circulation, are used up, and glycogen appears in the liver. If, however, the sugar in the blood rises above about 0·18 per cent. (the *renal sugar threshold*), glucose appears in the urine.

The occurrence of glycosuria depends upon two things: (1) the sugar content of the blood; (2) the renal sugar threshold. When glycosuria occurs it means that the blood sugar has risen to a value higher than the renal threshold (normally 180 mg./100 c.c.); consequently, glycosurias may be due to one of two classes of cause: first, to an abnormally high blood sugar; second, to an abnormally low kidney sugar threshold. The latter type, occurring naturally, is called *renal glycosuria*, and has no further interest for us here, apart from the importance of distinguishing between it and other glycosurias which are more serious.

(1) **Alimentary Glycosuria.** Hyperglycæmia may be induced by the administration of large quantities of glucose by the mouth. Glycosuria from this cause in normal people, however, is largely a figment of the imagination, since the body can dispose of the excess of glucose as fast as it is absorbed; even when 500 grammes of glucose is taken no sugar may be passed.* In certain individuals the power of utilising glucose may be deficient, so that prolonged hyperglycæmia and copious glycosuria may be caused by any over-indulgence in carbohydrate food. This condition is revealed by a glucose tolerance test.

(2) **Influence of the Nervous System.** The liver is richly supplied with sympathetic nerves, division of which does not seem to affect the glycogenic function.† Stimulation of them, however (after removal of the suprarenals), produces rapid reduction of the glycogen content of the liver, hyperglycæmia and glycosuria.

(3) **Adrenaline.** The injection of small amounts of adrenaline into the circulation has a similar effect to that of stimulation of the splanchnic nerves.

Asphyxia, the administration of anæsthetics, infective processes, pain, emotion, all cause depletion of liver glycogen, usually with a rise of blood sugar, and often with glycosuria. Most of these factors act by causing a discharge of impulses along the sympathetic nervous system, which include among their effects a liberation of adrenaline into the circulating blood. If the suprarenals are previously removed, these conditions cause a smaller rise of blood sugar and smaller loss of liver glycogen. If the liver is removed no rise of blood sugar is caused in these conditions.

---

* TAYLOR and HULTON. *J. Biol. Chem.*, 1916, **25**, 173; BENNETT and DODDS. *J. Physiol.*, 1925, **59**, Proc. lxiii.

† DONALD. *Amer. J. Physiol.*, 1931, **98**, 605.

Prolonged administration of adrenaline, as by subcutaneous administration, may give different results. This is because the Cori cycle becomes increasingly important, with the result that muscle glycogen is converted first into lactic acid and then into liver glycogen, which at the end may be raised.

(4) **Piqûre Glycosuria.** It was shown by Claude Bernard,* using rabbits, that puncture at a certain spot on the floor of the fourth ventricle was immediately followed by glycosuria. This "piqûre" glycosuria may last for twenty-four hours or more, and is accompanied by hyperglycæmia, which is more evident if the animal has previously been well fed.

It is due to the stimulation of a group of nerve cells from which impulses pass out by way of the splanchnic nerves to the suprarenals and liver. Hence if both splanchnic nerves have previously been divided the hyperglycæmia and glycosuria are much reduced or altogether lacking. The "centre" can be reflexly affected by stimulation of various afferent nerves, such as the vagus, sensory nerves, etc.

Effects somewhat similar are seen when the central nervous system is transected at the level of the superior corpora quadrigemina.† The same condition is seen on removal of the cerebral cortex,‡ or section of the hypothalamus,§ after which operations there is a sustained excitation of the whole sympathetic nervous system.

Piqûre or decerebration causes a large outpouring of sugar into the blood, even in fasting animals, and in these the amount of sugar released may be greater than can be accounted for from the observed reduction of glycogen in liver and muscles.‖ Indeed, in fasting animals, piqûre or decerebration may even produce an increase of liver glycogen, in spite of the large amount of sugar which is excreted. It would seem probable, therefore, that the consequences of piqûre may be three-fold : first, in the well-fed animal a discharge of liver glycogen in the form of excreted glucose ; second, transfer (often only transitorily) of muscle glycogen to the liver, by way of the Cori cycle ; third, in the fasting animal, in addition to the discharge of glycogen from the liver, processes of glyconeogenesis are accelerated in the liver.

(5) **Phlorrhizin Glycosuria.** Phlorrhizin is a glucoside of phloretin from the root cortex of the apple tree. When given suspended in olive oil or propylene glycol, by subcutaneous injection in doses of 1 g. per 10 kg. per day, it gives rise to intense polyuria and glycosuria. The glycosuria is not due to hyperglycæmia. The blood sugar is, in fact, diminished. Phlorrhizin glycosuria is due to an action on the kidney ; the condition resembles renal glycosuria, but the renal threshold is here quite abolished. If cannulæ be placed in the two ureters so as to collect the urine from each kidney separately, and a small dose of phlorrhizin be then injected into the left renal artery, the urine flowing from the left ureter will in two minutes be found to contain sugar, while the urine from the right kidney will not contain any sugar for another five or ten minutes. In the normal kidney, the glucose passing into the glomerular filtrate is all reabsorbed into the blood, in the proximal convoluted tubule, and so does not reach the urine (*v.* Chapter XLVII). In phlorrhizin poisoning the reabsorption fails, and so the glucose, together with considerable water reaches the urine. The action of phlorrhizin is due to its inhibiting the phosphorylation of glucose, which is an essential step in its absorption by the kidney tubules.

Phlorrhizin thus continually drains off sugar from the blood, and the various effects which it produces are consequences of this. As some of the effects closely resemble the phenomena of diabetes mellitus, it was once thought that the fundamental causes were similar, but this is not so.

Because of the lowered glucose content of the blood, the liver produces fresh sugar to take the place of that lost. At first the liver will convert most of its stored-up glycogen, but if a second dose of phlorrhizin be given when the liver may be regarded as nearly free from glycogen, glycosuria is produced as before, and the excretion of sugar can be continued indefinitely by repeated administration of the drug.

So long as sufficient carbohydrate food is given, the loss of sugar does not entail any increase in the destruction of the tissues ; but otherwise the waste of sugar has to be made good at the expense of material other than carbohydrate. The readiest source of the sugar excreted by a starved phlorrhizinised animal is protein, hence there is a connection, though no constant ratio, between the amounts of nitrogen and of sugar excreted. In different experiments this mean G : N ratio varies from 1·0 : 1 to 3·6 : 1.

---

* C. Bernard. "Leçons de Physiologie experimentale" (Paris), 1855, p. 296 "Système nerveux" (Paris), 1857, p. 397.

† J. Mellanby. *J. Physiol.*, 1919, **53,** 1.

‡ Bulatao and Cannon. *Amer. J. Physiol.*, 1925, **72,** 295.

§ Beattie, Brow and Long. *Proc. Roy. Soc.*, 1930, **106B,** 253.

‖ Donhoffer and Macleod, *Proc. Roy. Soc.*, 1932, **110B,** 141, 158 ; Evans, Tsai an Young. *J. Physiology*, 1931, **73,** 81.

Glucose administered to a phlorrhizinised dog is practically all recoverable from the urine.[*] The sugar excreted in starvation must be derived largely from the amino-acids of tissue protein. It was shown by Lusk, Embden, and Dakin that the amino-acids now called anti-ketogenic (p. 928) yield glucose when administered to a phlorrhizinised animal. On the other hand, the ketogenic amino-acids give no increase in the output of sugar, but are converted into aceto-acetic acid. The administration of fats to a phlorrhizinised dog is also without effect on the sugar excretion.

Another effect of continued administration of phlorrhizin is fat infiltration of the liver, a result of the carbohydrate starvation. The liver seems to be able to act as a storehouse either of fat or of carbohydrate, so that there is an inverse ratio between the amount of glycogen and the amount of fat stored up in the liver at any given time.

For Reference

McKee and Hawkins (1945). Phlorhizin Glucosuria. *Physiol. Rev.*, **25**, 255.
Nash (1927). Phlorrhizin Diabetes. *Physiol. Rev.*, **7**, 385.

Ketosis. We have seen (p. 947) that, when carbohydrate oxidation is suppressed, the last stages in the oxidation of fats make default, so that the acetone bodies are formed in the liver, accumulate in large quantities in the blood (*ketonœmia*), and are excreted in the urine (*ketonuria*). The appearance of these keto-acids and acetone is spoken of as *ketosis*, and sugar is said to be *antiketogenic*. There is a relationship between the amounts of carbohydrate and fats which are necessary to prevent ketosis. It has been shown that no ketosis occurs when one molecule of sugar is oxidised in the body for each molecule of aceto-acetic acid which may be formed from the fats (or a portion of the proteins) of the food. When the proportion is only 1 to 2, there is always some ketosis, and when the proportion of potential aceto-acetic acid to glucose exceeds 2, the whole excess is excreted as ketone bodies. This relationship is of practical importance in the dieting of diabetics.

Not only does the excretion of ketone bodies involve a loss of energy to the body, but these organic acids require bases for their neutralisation. Their excretion therefore robs the body of available base (acidosis). When the supply of these fails the acids are excreted in combination with ammonia. If the condition of carbohydrate starvation be continued, this mechanism of neutralisation is insufficient, and dyspnœa and coma ensue as the death of the animal approaches. The coma is accompanied by failure of the brain to utilise oxygen.[†]

(6) **Pancreatic Diabetes.** Von Mering and Minkovski[‡] found that pancreatectomy (excision of the pancreas) gives rise to a rapidly fatal diabetes, similar to the severest cases of diabetes in man. The operation is followed almost at once by hyperglycæmia, which may exceed 1,000 mg./100 c.c., and by the appearance of a large amount of sugar, 5 to 10 per cent., in the urine. The glycosuria persists ; if glucose is given it is all excreted ; there is polyuria and thirst, the urinary nitrogen is raised, but there is not,[§] as formerly supposed, a constant G/N ratio during fasting. The respiratory quotient is low (about 0·7) ; acetone bodies are excreted ; liver glycogen almost disappears, and muscle glycogen is much reduced, but the glycogen in the heart is increased ; the glucose uptake of some of the tissues is reduced, as is shown by the slight difference in glucose content between arterial and venous bloods ; the animal rapidly wastes, and finally dies in diabetic coma at the end of a week or so. The condition of diabetes observed under these circumstances has nothing to do with the absence of the pancreatic juice from the intestines, since ligature of the ducts of the pancreas, or grafting the pancreas beneath the skin, does not give rise to diabetes, although these procedures are followed by atrophy of the secretory part of the gland.[||]

---

[*] Reilly, Nolan and Lusk. *Amer. J. Physiol.*, 1898, **1**, 395.
[†] Kety *et al.* *J. Clin. Inv.*, 1948, **27**, 500.
[‡] Von Mering and Minkovski. *Arch. f. exp. Path.*, 1889, **26**, 371.
[§] Soskin. *J. Nutr.*, 1930, **3**, 99.
[||] Ssobolew. *Virchow's Arch.*, 1902, **168**, 91.

We thus see that the pancreas has a two-fold function: the secretion of the pancreatic juice, and an influence, the absence of which is followed by diabetes. Corresponding with this two-fold function, two kinds of structures are present in the gland, the secreting acini and the islets of Langerhans. These latter, though arising in connection with the ducts, are solid masses of cells and have no communication with the ducts. The stainable islet cells may be divided into two varieties,[*][†] which have been given the name of α and β cells, according as their granules are fixed respectively by alcoholic or watery solutions. These cells undergo no alterations when the gland is excited to secrete [‡] pancreatic juice (Fig. 596). On the other hand, if four-fifths of the pancreas be removed, the remaining part may gradually become inadequate to prevent diabetes; and when under these circumstances diabetes supervenes, the granules disappear from the β cells; also, the insulin content of the remnant falls to a very low level.[§] Changes have also been alleged in the islets of Langerhans in fatal cases of diabetes in man.

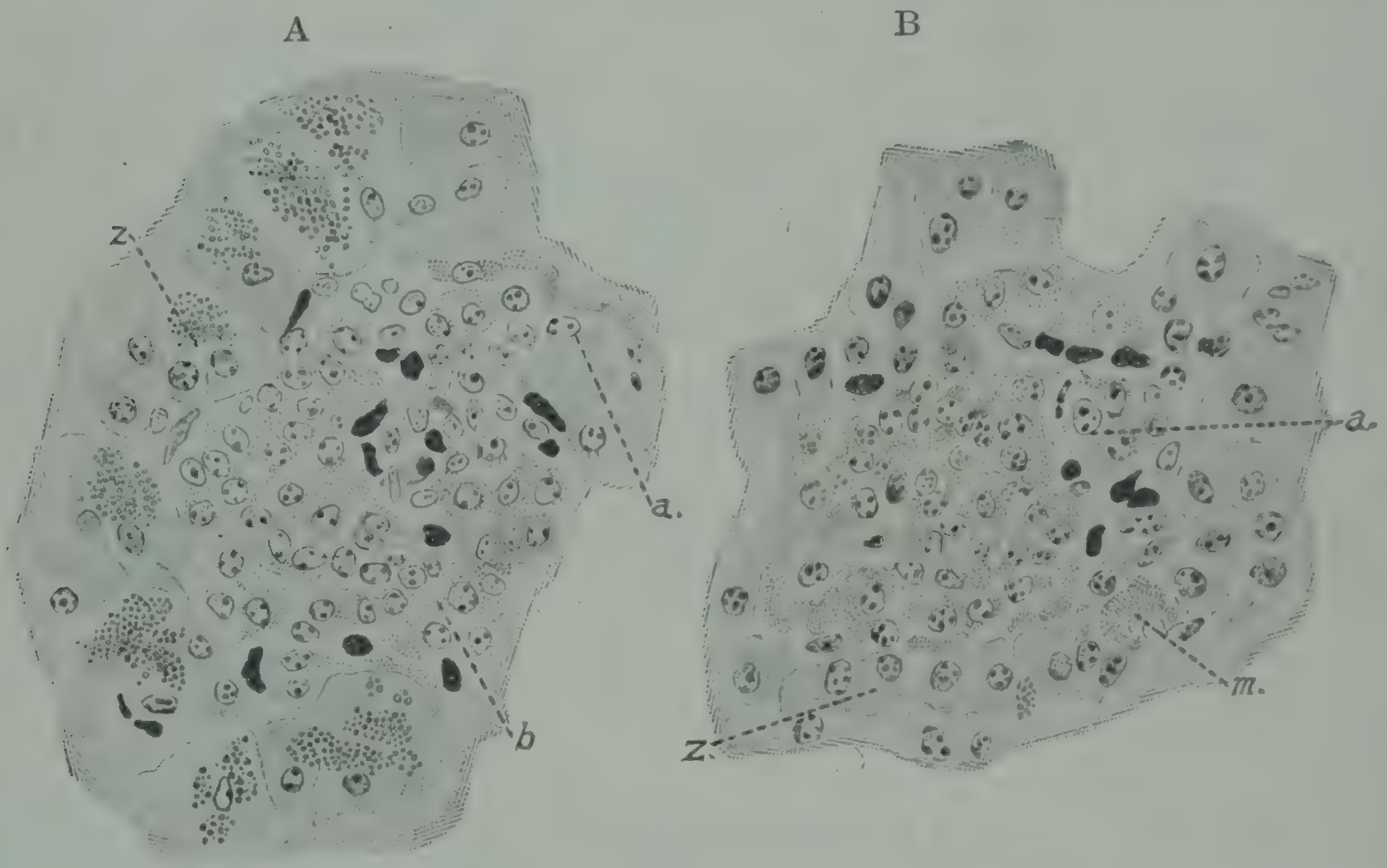


FIG. 596. (A) and (B) show an Islet with the surrounding Tissue in a resting Gland (A) and after Exhaustion with Secretin (B).

In (A) the secreting acini (z) are charged with zymogen granules. In (B) these have entirely disappeared. On the other hand no change is noticeable in the cells of the islet. In the latter the granular cells (m) are the β cells, and the clear hyaline cells are the α cells. (HOMANS.)

If large doses of alloxan (200 mg./kg.) are injected into the circulation there is at first a rise in blood sugar, then a fall, and finally, after eighteen to twenty-four hours, a permanent hyperglycæmia, and the animal becomes very ill. The initial rise of blood sugar is a sympathetic response, the succeeding fall, at six to twelve hours, is due to liberation of pre-formed insulin from the β-cells.[||] These are damaged within a few minutes of the injection, and do not recover; hence the permanent hyperglycæmia; it is found that the β-cells in the islets of Langerhans have degenerated, and a state of diabetes, equivalent to a partial pancreatectomy, results.[¶]

There are two classical theories of the nature of the diabetic state; these attribute the condition respectively to non- (or under-) utilisation of glucose by the tissues or to over-production of glucose by the liver. Many of the facts fit in with either theory. The earlier view of under-utilisation, which has since lost ground, would suppose that the supply of sugar to the body by the liver remains unaltered, but the demand by the tissues is reduced or abolished, with consequent hyperglycæmia and glycosuria. Increased protein and fat metabolism is necessitated to provide the energy, but also leads to increased sugar formation. The opposing view considers that the liver regulating mechanism is "set" to a high blood-sugar level, and accordingly the liberation and formation of sugar by the liver is accelerated; hence the increased protein, and fat, metabolism; the tissues do not demand, and cannot use, the increased amount of sugar, and the excess is excreted.

---

[*] LANE. *Amer. J. Anat.*, 1907, **7**, 409.

[†] BENSLEY. *Amer. J. Anat.*, 1911, **12**, 297; *Harvey Lects.*, 1914–15, 250.

[‡] HOMANS. *Proc. Roy. Soc.*, 1913, **86B**, 73.

[§] BELL *et al.* *J. Physiol.*, 1942, **101**, 11.

[||] HUGHES, WARE and YOUNG. *Lancet*, 1944, (i), 148.

[¶] SHAW DUNN *et al.* *J. Physiol.*, 1944, **103**, 233.

Uncontrovertible evidence that diabetic tissues can still utilise glucose was provided by the observation that after removal of the liver, the blood glucose fell at an even more rapid rate in diabetic (depancreatised) than in non-diabetic dogs, and ultimately both sorts of animal ended in hypoglycæmic convulsions.* Further, it has been shown that in order to maintain the blood sugar in the liverless dog, glucose must be infused into the blood stream at the rate of about 0·25 g./kg./hr., and a similar rate is needed to keep up the sugar level in the *diabetic* liverless ("eviscerated") dog.

Whichever theory is invoked, there seems to be little doubt that much of the sugar excreted by fasting diabetics arises from proteins. At one time it was generally supposed that the G/N ratio was constant, and that this pointed to protein as the sole source of the neogenetic glucose, but it now seems clear that the observations were strained to fit this hypothesis, and that quantitative data are both inadequate and of equivocal significance.

Since it must be conceded that diabetic tissues do utilise glucose, it is doubtful whether protein alone could provide all the sugar that is excreted, and this means that gluconeogenesis from fat is likely. Against this has been urged that feeding on fat does not raise the sugar excretion. But calculations based on sugar utilisation and excretion in depancreatised and/or hepatectomised dogs gave strong indications that in the fasting diabetic the production of sugar from fat proceeds at the maximal rate,† so that giving fat in the diet could not augment it.

Some direct experiments give support to this view, since liver perfusions and liver tissue-slice experiments showed a conversion of fat, or fatty acids, into glucose.‡

The low R.Q. in diabetes and the absence of a rise if carbohydrate is ingested, have in the past been taken to mean that carbohydrate oxidation is greatly lowered, and that combustion of fat has taken its place. But, in fact, its meaning is doubtful, since all metabolic processes are deranged, and the R.Q. is of such a composite nature as to preclude any firm conclusions.

Although the diabetic tissues continue to utilise, and probably to oxidise, glucose, they do not seem to do so as easily as normal tissues can. Glycogen formation in muscle is sluggish and the rate at which glucose is converted into fatty acids is lowered. The rate of glucose uptake by the tissues varies with the glucose concentration of the blood, and in the diabetic this level is greatly raised. At this raised level the rate of utilisation of glucose by the diabetic is about the same as that of the normal tissues with normal blood sugar level. The fundamental change in the diabetic state therefore seems to be that both the liver and the remaining tissues are "set" to a high sugar threshold; the liver continues to produce sugar even when the blood sugar is already high; the tissues utilise sugar, even at this high level, no more readily than normal tissues would do at normal levels; but at these high levels the kidney threshold is passed and sugar is excreted. Hence liver glycogen is very low, muscle glycogen is low, glyconeogenesis is accelerated, and fat metabolism increased; fat accumulates in the liver, and there is abundant fat in the plasma (lipæmia). Administration of glucose to the animal causes an almost corresponding increase in the amount of glucose excreted in the urine; this is probably due to the fact that sugar is already being formed by gluconeogenesis at a rate too fast for the tissues to utilise it, together with the failure of the liver and muscles to store glycogen. With increased production and diminished storage of carbohydrate, the blood is flooded with sugar, which escapes by the kidneys, while the accelerated breakdown of fat leads to the formation of ketone bodies in large amounts. These ketones can probably be used, up to a point, by the tissues, *in lieu* of glucose. But beyond a certain point (about 2·5 g. of fat per kg. per day at rest), the ketones cannot be so utilised, and are excreted in the urine.

---

* MANN and MAGATH. *Arch. Int. Med.*, 1923, **31**, 797.

† YOUNG. *J. Physiol.*, 1936, **87**, 11P.

‡ BURN and MARKS, *J. Physiol.*, 1926, **61**, 497; BLIXENKRONE-MØLLER, *Z. f. Physiol. Chem.*, 1938, **252**, 137.

INSULIN

Many facts led physiologists to believe that the pancreas produced, in the islets of Langerhans, a hormone necessary for the normal metabolism of carbohydrates, but it was not till 1922 that Banting and Best[*] succeeded in extracting the hormone from normal pancreas, and in showing that, by its injection, it was possible to obviate the effects of removal of the pancreas in dogs, and by regular dosage to effect a restoration to health in cases of diabetes in man.

In order to prepare insulin, the fresh pancreas is finely minced and extracted with several changes of alcohol which, diluted by the water contained in the gland, is evaporated to a small bulk and poured into absolute alcohol. A white precipitate is produced which represents the crude insulin. The yield of insulin is increased by extracting in the first instance with alcohol containing sodium bicarbonate. Insulin forms an insoluble compound with picric acid, which is decomposed by weak hydrochloric acid with the formation of an hydrochloride. These properties are made use of to effect a further concentration of the active principle.[†]

The insulin so prepared is a white powder, obtainable in crystalline form,[‡] easily soluble in water and in dilute alcohol. As usually prepared, crystalline insulin contains zinc.[§] It does not dialyse, will withstand boiling for a short time, is stable in acid solutions, and undergoes slow oxidation in alkaline solution. It is a protein of molecular weight 35,000, contains sulphur, and is destroyed by pepsin and trypsin, so that it is ineffective if given by mouth. 0·2 mg. injected into a normal rabbit causes a rapid diminution of the blood sugar, and finally the symptoms of hypoglycæmia.

In has long been known that with some preparations of insulin a transitory rise of blood sugar precedes the fall. This is due to the presence of an insulin-like component which accelerates gylcogenolysis in the liver. The distribution of the substance, and its persistence in the pancreas of rabbits deprived of β-cells by alloxan suggests that it is present in the α-cells.[||]

The strength of insulin preparations must be carefully controlled by biological standardisation. The method is based on the reduction of the blood-sugar level of animals by the sample. Up to a certain dose, of about 1 unit/kg., the *rate of fall* of blood sugar increases with the dose, and the per cent. fall of sugar is proportional to log. dose. Larger doses than 1 unit per kg. produce longer lasting falls of blood sugar. When the true sugar reaches a particular low level, probably near to zero, convulsions occur. A dose of about 2 units per kg. will in four hours lower the blood sugar level of a fasting rabbit to

[*] Banting and Best, *J. Lab. Clin. Med.*, 1922, **7**, 251, 464; Banting, Best and Macleod, *Amer. J. Physiol.*, 1922, **59**, 479.

[†] Dudley, *Bioch. J.*, 1923, **17**, 376; 1924, **18**, 665; Dodds and Dickens, *Lancet*, 1924 (i), 330; *Brit. J. Exp. Path.*, 1924, **5**, 115.

[‡] Harington and Scott. *Bioch. J.*, 1929, **23**, 384.

[§] Scott. *Biochem. J.*, 1934, **28**, 1592.

[||] Sunderland and de Duve, *J. Biol. Chem.*, 1948, **175**, 663.

the convulsive level. The present method is to determine the fall of blood sugar brought about in five hours in a fasting rabbit by subcutaneous injection of known amounts of the solution to be tested, as compared with that induced by the international standard preparation, of which 1 mg. contains 22 units. The purest crystalline insulin has a potency of about 23 units per mg.

Human pancreas contains about 1·7 units of insulin per g., and most mammalian pancreas gives similar figures. In the islet tissue it has been calculated that there is about 150 units per g. Presumably it is liberated at an appropriate rate into the blood, though it has not been demonstrated there, and then is taken up by the tissues.

If insulin be injected into a pancreatectomised animal the excretion both of sugar and of ketone bodies ceases, arterio-venous glucose difference rises, showing that the animal is utilising sugar; and on the administration of carbohydrates glycogen accumulates in the muscles and liver. The animal increases at the same time in weight and strength, and can be maintained alive for its normal span. It is necessary, however, to inject insulin subcutaneously two or three times a day, since it is gradually used up in the body.

One of the first questions concerning the action of insulin is the fate of the blood sugar which so rapidly disappears when insulin is given. In the normal animal the tissue carbohydrate and blood sugar may be regarded as derived either directly from ingested carbohydrate, or indirectly from glycogen formed from blood sugar, or from the breakdown of protein, and to some extent also from fat. Whether carbohydrate is ingested or not, its presence and continual metabolism in the tissues are essential to life.

In the diabetic subject, it seems quite clear that the effect of administering suitable amounts of insulin is to restore the carbohydrate metabolism, and hence metabolism in general, to normal; in other words, to accelerate the uptake of carbohydrate, to check glyconeogenesis, and to allow the liver and muscles to carry a store of glycogen; hence, to lower the blood sugar; because the carbohydrate metabolism is restored, and because new formation of carbohydrate is reduced, ketone bodies speedily disappear from the blood, and therefore from the urine and breath. It also accelerates the rate of conversion of glucose to fatty acids, lowers the blood amino-acids, the inorganic PO_4''' of plasma, and blood potassium. It is evident that, in the diabetic, insulin produces all these effects by supplementing the inadequate action of the pancreas.

How insulin acts is not exactly known. In the normal animal or man, there is already an efficiently functioning pancreas, so that by administering insulin to them we shall get abnormal effects due to the presence of an excess of that substance. We shall consider its actions later.

ENDOCRINE INFLUENCES

In addition to the effects of insulin, carbohydrate metabolism is affected by internal secretions set free into the blood by the adrenal medulla (adrenaline), the adrenal cortex, the thyroid and the anterior pituitary. The effects of adrenaline have already been considered.

Influence of Anterior Pituitary Hormone (see also Chap. L.). In diseases associated with enlargement of the anterior lobe of the pituitary body (*e.g.* acromegaly and pituitary basophilism) there is in 20 or 30 per cent. of cases, hyperglycæmia and glycosuria.

The work of Houssay and of Young has shown that the anterior pituitary body produces hormones which have fundamental importance in the regulation of carbohydrate metabolism.

Hypophysectomy. After hypophysectomy (removal of the pituitary body), no serious effects are seen if the animals are well fed. Their nitrogen excretion is much below normal, which means that glyconeogenesis is slowed down, the total metabolism is low and there is a low blood sugar. Carbohydrate oxidation is augmented.

On fasting, or on any treatment that lowers the glycogen stores of the body, the blood sugar falls to a low level, and hypoglycæmic seizures ensue. The animals when fasting are demonstrably much (often 30 times) more sensitive to the action of insulin than are normal animals,* and less responsive to adrenaline.

One of the results of hypophysectomy is a reduced activity of other endocrine organs, *e.g.* thyroid and adrenals, and some of its effects may be secondary to such alterations.

Combined Pancreatectomy and Hypophysectomy. The results of removal of both pancreas and pituitary are most striking †: the state of diabetes, which usually follows the extirpation of the pancreas, is held in abeyance. Animals subject to the double extirpation are often called " Houssay " animals : if fed moderately, they show practically normal blood sugars, little if any glucose excretion, no ketosis, normal glucose tolerance, and normal glycogen

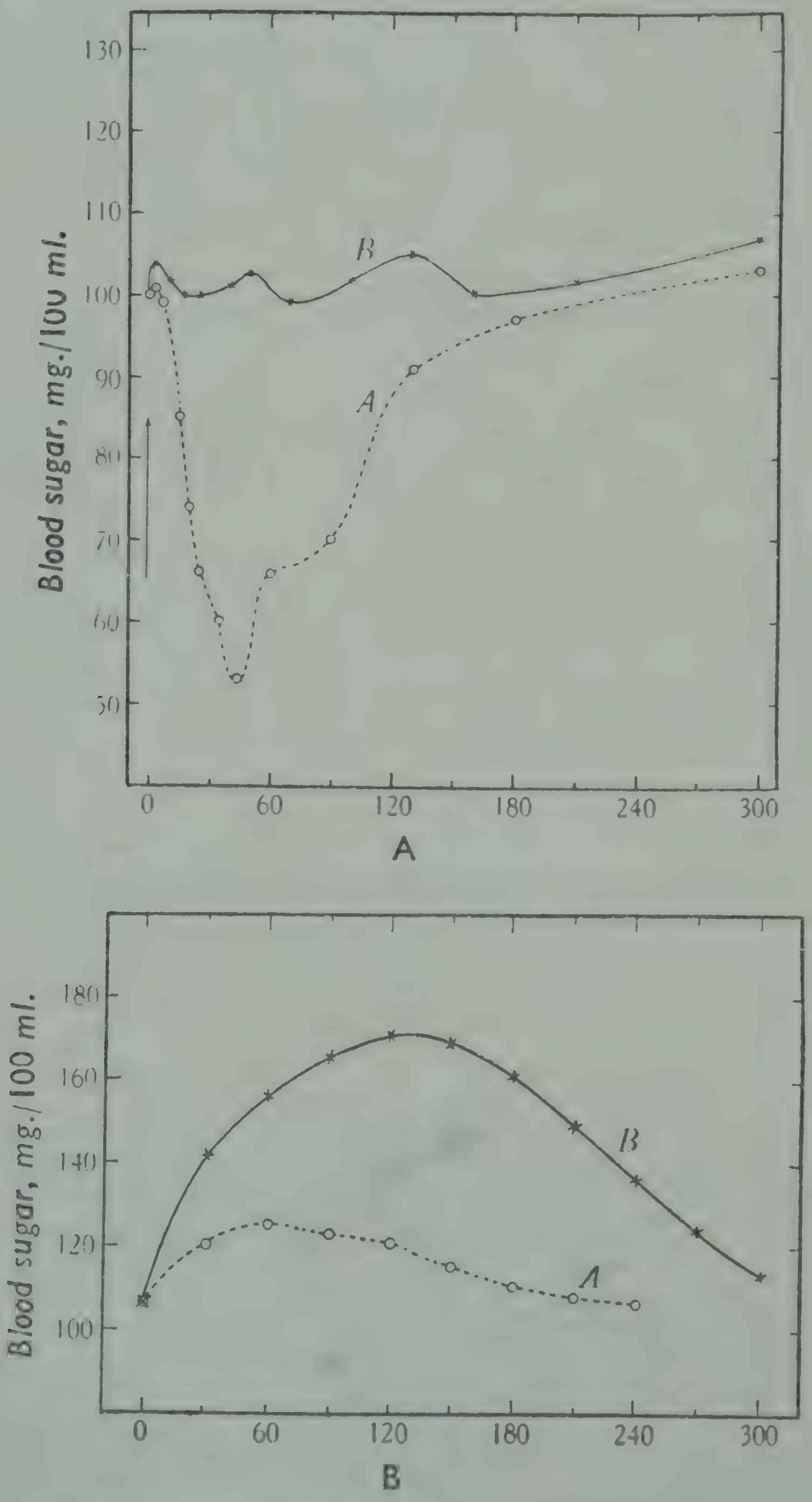


FIG. 597. Effect of anterior pituitary extract on the sensitivity of rabbits to insulin (Graph A) and adrenaline (Graph B).

A. Normal rabbits. B. Rabbits previously treated with anterior pituitary extract. (YOUNG, *Bioch. J.*, 1938, **32**, 1521.)

---

* CORKILL, MARKS and WHITE. *J. Physiol.*, 1933, **80**, 193.

† HOUSSAY and BIASOTTI. *Endocrinol.*, 1931, **15**, 511.

stores in liver and muscles. Their nitrogen excretion is about normal, and they survive for weeks or months, but are subject to hypoglycæmic crises if subjected to fasting. It seems evident that their carbohydrate metabolism is unstable and is easily tipped in the direction of the hypoglycæmic state.

Injection of Pituitary Extracts. The number of hormones now claimed to be formed by the pituitary anterior lobe is considerable, and there is some doubt about the individual existence of some of them. But by the injection of extracts of the anterior pituitary, prepared in various ways, a diabetogenic effect is produced. This effect comprises an accelerated rate of gluconeogenesis in the liver, an opposition of the action of insulin, hyperglycæmia, etc.

The Diabetogenic Hormone. By repeated intraperitoneal injections of such extracts of the anterior pituitary into dogs, symptoms resembling those of diabetes develop after a few days : these symptoms comprise reduction of the insulin content of the pancreas, depletion of liver glycogen, hyperglycæmia, glycosuria, ketonuria, high nitrogen output, etc. Extirpation of the liver prevents these results. Given to "Houssay" animals, the injections cause a revealing of all the latent diabetic symptoms.

Usually the symptoms resulting from a few injections subside soon after discontinuance of the dosage. But after repeated dosage there results a permanent condition of diabetes (metahypophyseal diabetes).* In animals thus rendered permanently glycosuric, the islets of Langerhans were found to have undergone degenerative changes.† These are presumably due to overwork of the β-cells, which, during the injections, secrete insulin at ten to twenty times the normal rate ; if enough insulin is given with the pituitary extract, these changes are obviated.‡

The most usual explanation of the facts is that the hormones of the pituitary exert an influence in favouring glyconeogenesis, in opposition to insulin, which appears to check it. Also, they counteract the peripheral action of insulin on the tissues, and hence reduce the uptake of sugar from the blood. After hypophysectomy, the nitrogen output is reduced, while administration of pituitary extracts raises it. We may thus suppose that under ordinary conditions, in which the pituitary hormones are present, a good deal of glyconeogenesis is going on, so that, even in starvation, glucose is still being added to the blood by the liver. The administration of insulin to the fasting animal may somewhat tax this supply of carbohydrate, but after hypophysectomy the endogenous supply of carbohydrate is much diminished, while demands by the tissues are not reduced, so that the animals become very much more sensitive to the hypoglycæmic action of insulin. With the action of adrenaline it is the reverse of this—carbohydrate is most available when liver glycogen is high, and labile, so that hypophysectomy reduces the response to adrenaline, and the administration of pituitary extracts enhances it. When both pancreas and pituitary are removed the animal's carbohydrate metabolism somewhat resembles that of the normal state, provided it is well fed, and thus given a supply of ready-made carbohydrate to draw upon. Then, although glyconeogenesis is reduced by loss of the pituitary, there is at the same time less tendency to hypoglycæmia because the pancreas is also lacking, and hence there is sufficient carbohydrate to meet most ordinary needs.

The amount of carbohydrate consumed in the normal body is subject to wide variations, dependent on the supply of carbohydrate in the diet. When the supply is restricted, it is endogenously produced from protein and probably also from fat, so that the irreducible minimum of blood sugar is available as long as life lasts. Determination of the glucose-tolerance curves shows that on low carbohydrate diets the tolerance is reduced and on carbohydrate-rich diets it is augmented, *i.e.* the more carbohydrate in the diet the more readily can administered glucose be disposed of * (Fig. 598). Hypophysectomy abolishes this effect of diet. Low carbohydrate diet also reduces the response to insulin. All these facts taken together suggest that the pituitary hormones are secreted in greater amount when the diet is poor in carbohydrate and *vice versâ*.† After prolonged total fasting the capacity to utilise ingested sugar is much impaired, so that when sugar is given it may at first be excreted unchanged (hunger diabetes). The lost ability to use glucose is not due to inability to store glycogen in the liver ‡ ; it is possibly due to the fact that glyconeogenesis has been going on intensively, the new-formed glycogen being at once converted into glucose, the usage of which is reduced. When glucose is given this floods the organism until such time as the excessive gluconeogenesis can be checked and the rate of usage restored.

---

* YOUNG, *Lancet*, 1937, ii, 372 ; *Biochem. J.*, 1938, **32**, 513, 524.

† RICHARDSON and YOUNG. *Lancet*, 1938, *i.*, 1098.

‡ BEST *et al.* *J. Physiol.*, 1942, **101**, 17.

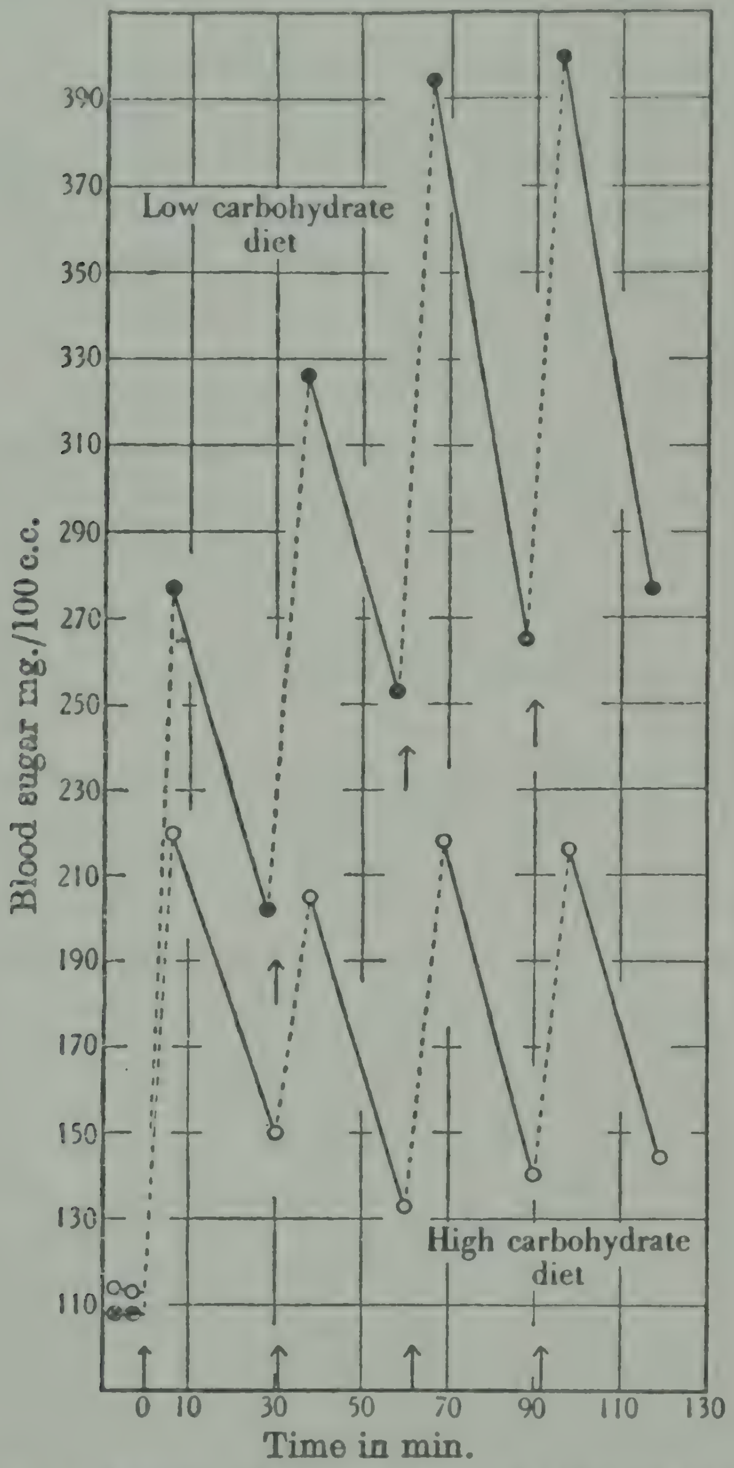


FIG. 598. Response of the same rabbit on high and low carbohydrate diets to half-hourly intravenous injections of 1 gramme glucose. (HIMSWORTH and SCOTT, *J. Physiol.*, 1938, **91**, 447.)

Influence of Adrenal Cortex. The adrenal body exerts an effect on carbohydrate metabolism not only through its medulla, which liberates adrenaline into the blood, but also through the cortex. It is well recognised that after removal of the suprarenals the absorption of monosaccharides from the intestine is retarded, and the glycogen stores of the body somewhat reduced even if full diet is maintained ; in fasting the glycogen stores are rapidly depleted, while the blood sugar is lowered and the blood lactate raised.

---

* HIMSWORTH. *Clin. Sci.*, 1935, **2**, 67.
† HIMSWORTH and SCOTT. *J. Physiol.*, 1938, **91**, 447.
‡ DANN and CHAMBERS. *J. Biol. Chem.*, 1932, **95**, 413.

Further study shows that the effects of removal of the adrenals somewhat resemble those of hypophysectomy. Thus there is the same tendency to hypoglycæmia on fasting, the lowered glycogen reserves, especially on fasting, the sensitivity towards insulin, the raised oxidation of carbohydrate, with raised R.Q., the depressed glyconeogenesis, and, what is even more important, the same amelioration of diabetic symptoms in the pancreatectomised animal. The diabetogenic action of pituitary extracts, moreover, is reduced by adrenalectomy. This is especially well seen in the partially depancreatised rat, which is very sensitive to anterior pituitary hormones (Fig. 599).

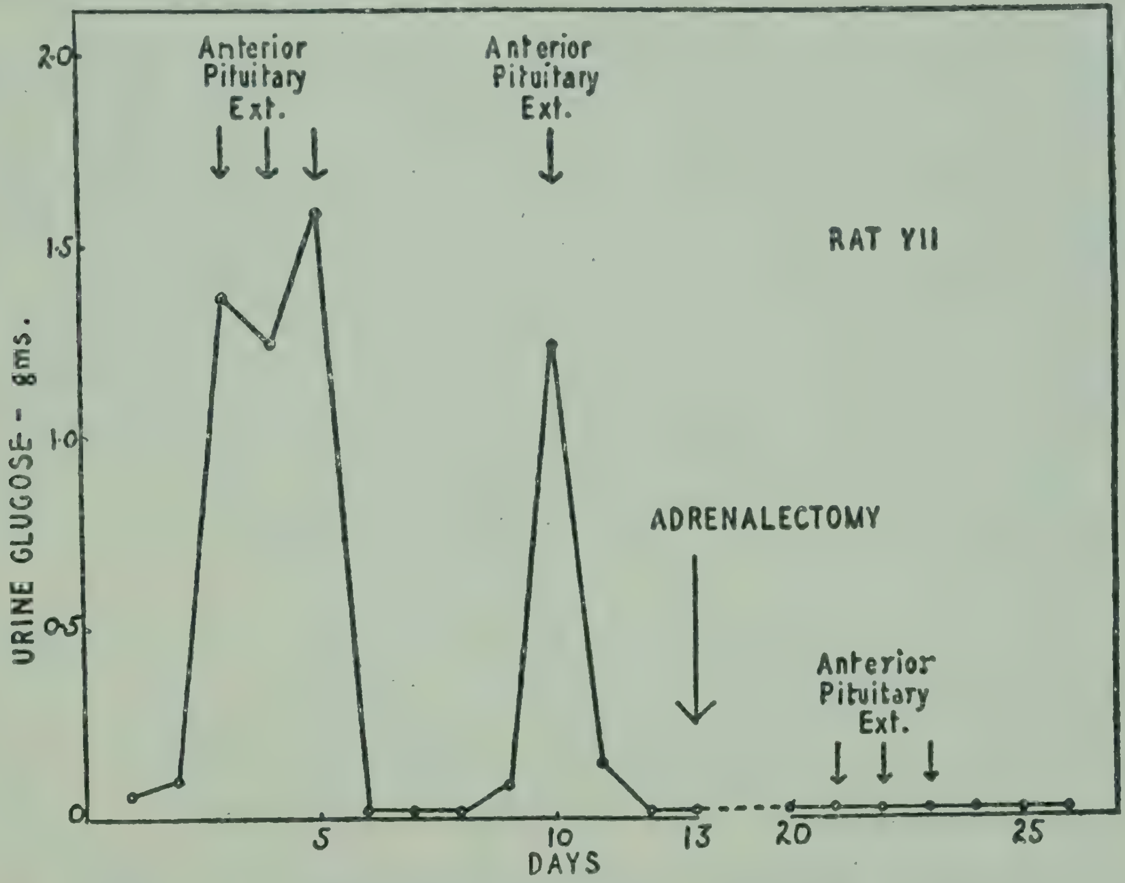


FIG. 599. Effect of adrenalectomy on the glycosuric response to anterior pituitary extract in the partially depancreatised rat. (LONG. *Harvey Lectures*, 1937.)

When extracts of the adrenal cortex, or the cortical hormones (the ones with O or OH at position 11) are injected into adrenalectomised animals, the effects of the adrenalectomy are reversed. If the adrenal cortical extract is given to fasting hypophysectomised rats there is a rise in the blood sugar and in the store of glycogen, as the following summary shows :—

MEAN CARBOHYDRATE LEVELS OF FASTED RATS INJECTED WITH CORTICAL HORMONE *

Treatment	Muscle Glycogen. m.g./100g.	Liver Glycogen. m.g./100g.	Blood Glucose. m.g./100g.
Normals, untreated	521	188	84
Hypophysectomised, untreated .	320	94	52
Hypophysectomised, given 0·03 mg. Kendall's Adrenal Comp. B every hour for 24 hours .	450	1,520	107

It may reasonably be concluded that the adrenal corticosteroids mainly influence carbohydrate metabolism through their action on glycogen storage, particularly and primarily in the liver. This action is due to reduced oxidation of carbohydrate, reduced influence of insulin, and increased glyconeogenesis.

It is probable that some of the effects connected with the pituitary body are actually effected through a stimulating effect which this has, through one of its hormones, on the adrenal. After hypophysectomy, the adrenal lacks this stimulus to produce or liberate its cortical hormones, and the reduction of glyconeogenesis is partly due to this. Not all of the effects are likely to be due to this, however, since anterior pituitary extracts are still diabetogenic, though less so, after adrenalectomy. It is probably expedient, therefore, to assume that the pituitary hormones affect carbohydrate metabolism in two ways; first by their stimulating effect on the adrenal, and secondly by a direct effect in stimulating glyconeogenesis in the liver.

---

* LONG and KATZIN. *Proc. Soc. exp. Biol.*, 1938, **38**, 516 ; KATZIN and LONG. *Amer. J. Physiol.*, 1938, **123**, 113.

Influence of the Thyroid. The thyroid hormone (Ch. L.) has a powerful effect in increasing general metabolism, and so it is not surprising that increase or diminution of the hormone respectively stimulates or depresses carbohydrate metabolism. Most of the effects are probably a reflection of this action and do not represent a primary and specific action of the thyroid on carbohydrate metabolism.

Adrenal Cortex, Anterior Pituituary and Insulin. Fresh light has been thrown on the mode of action of insulin by study of its relation to hexokinase action. This enzyme catalyses the reaction by which glucose is converted into glucose-6-phosphate, which, as we have seen, is the first step in the formation of glycogen from glucose and in the oxidation or glycolysis of glucose. Hexokinase action has been shown * to be inhibited *in vitro* by anterior pituitary extracts, and the inhibition is accentuated if adrenal cortical extracts are also added. If insulin is added to such an inhibited enzyme system, the inhibition is removed and the activity restored. Insulin also enhances the hexokinase activity of diabetic muscle, and of muscle of rats previously injected with anterior pituitary extracts, but does *not* increase the hexokinase activity of normal muscle.

Although some discrepancies have still to be cleared up, this demonstration of such a fundamental action of insulin may well prove to be the key to all its physiological properties, since, by overcoming the inhibitory action of the pituitary-adrenal hormones it would virtually accelerate the formation of glycogen in liver and muscles, and thus " set " their regulating mechanisms to function at lowered blood-sugar levels. Reduced glyconeogenesis, increased conversion of glucose to fatty acids, and reduction of ketosis would probably be secondary consequences.

Nature of Insulin Action. We have seen that the blood-sugar level is to a large extent self-regulating, owing to its effect on the liver, which, when " set " to a particular level, will liberate or absorb glucose in such a way as to maintain that level. The setting of the level is dependent on the influence of various hormones, pituitary and adrenocortical hormones tending to " set " the liver to maintain a high blood sugar, and hence encouraging both glyconeogenesis and glycogenolysis, insulin tending to " set " the regulator to a low blood sugar, and hence checking both glyconeogenesis and glycogenolysis in the liver. Insulin, as Soskin has expressed it, " plays the part of a catalyst or activator in a process which can proceed at a slower rate in its absence. More specifically, it permits rates of carbohydrate utilisation at low blood-sugar levels which in its absence would require abnormally high blood-sugar levels."

The insulin liberated from the pancreas is mainly retained in the liver, and the influence of insulin on the whole animal is chiefly on the liver. This is shown by a calculation by Bouckaert and de Duve of glucose retentions by liver and extrahepatic tissues in the dog. The calculation expresses the glucose uptake, when blood sugar is kept at normal levels, in g./kg. body weight in eighty minutes (negative figures mean *release* of glucose into blood):

	Hepatic Retention.	Peripheral Tissue Retention.
(1) With normal amount of insulin .	− 0·21	+ 0·21
(2) With no insulin	− 0·69	+ 0·09
Difference = retention due to insulin.	+ 0·48	+ 0·12

---

* Colowick, Cori and Slein. *J. Biol. Chem.*, 1947, **168**, 583.

Whether carbohydrate oxidation is accelerated by insulin has been contested, but the balance of evidence indicates that it is. The total calorie production shows a slight rise, but the rise of R.Q. after giving insulin and glucose is not to be accepted without reserve, because conversion of glucose to fatty acid would also give a rise. If insulin is given to a normal individual, we get the effects of excess of it, and if at the same time the blood sugar is allowed to fall, as it normally would owing to increased uptake by the tissues, we find that liver glycogen is rapidly *reduced*. This is because glycogenolysis is increased owing to liberation of adrenaline, and at the same time glyconeogenesis is checked. But in normal endocrine balance insulin leads to glycogen deposition in both liver and muscle.

Best, Hoet and Marks * showed that when the blood sugar of the eviscerated spinal cat was kept at a high level for two to five hours by intravenous injection of glucose, there was no demonstrable change in the glycogen content of the skeletal muscles. When insulin was also given, there was an accelerated rate of sugar disappearance, and an *increase* in the glycogen content of the muscles, irrespective of whether actual hypoglycæmia resulted or not, and this deposition of the glycogen was enough to account for from 40 to 50 per cent. of the sugar which had been lost from the blood. It may be concluded, then, that under the conditions of these experiments, much of the lost sugar was recoverable as glycogen.

The Normal Regulation of Blood Sugar. Large amounts of carbohydrate can be dealt with by the normal individual without any appreciable loss of sugar by way of the urine ; on the other hand, owing to glyconeogenesis, abstinence from carbohydrates for quite long periods does not lead to hypoglycæmia. Now in the treatment of diabetes it is necessary to secure a proper balance between the amounts of insulin administered and carbohydrate ingested, as otherwise there will be either an excretion of glucose in the urine, or else the unpleasant effects of hypoglycæmia. In the normal state, since neither of these effects ensue, there must therefore be some means by which the rate of uptake of carbohydrate by the tissues corresponds with the amount of carbohydrate to be dealt with.

Up to a point this regulation is automatic, the liver adjusting its output or uptake of glucose, and the other tissues altering their rate of uptake according to the sugar concentration in the blood plasma. But there is evidence that the rate of manufacture, and of release, of insulin, is also controlled by the blood sugar level, *i.e.*, according to the need for it : fasting, feeding with fat, or administration of insulin for instance, all lead to a fall in the insulin content of the pancreas,† presumably because the rate of its formation is slowed down.

Apparently, when carbohydrate is taken by the mouth, there is an accompanying liberation of insulin to deal with it ; it is often seen that in a blood-sugar curve the blood sugar falls very rapidly after its initial rise, and reaches a level slightly below that of the blood in the fasting state (Fig. 589). The rapid fall is not because absorption of glucose from the alimentary canal is at an end ; in the diabetic subject the rapid fall is not seen. The fall in the normal subject without doubt occurs because by that time the insulin liberation has been raised sufficiently to reduce the blood sugar in spite of the continuing sugar absorption. So if a second 50 g. of glucose is given to a normal subject as soon as the blood sugar has reached its low level, there is little or no further rise in the blood sugar. It has even been shown that a transfusion into a diabetic subject of blood taken from a normal individual during the absorption of glucose will lead to a fall in the blood sugar of the patient, just as though insulin had been given.

It is uncertain how this regulation of insulin output by the pancreas is effected. Since it appears to be unaffected by denervation of the pancreas, it would seem to be independent of any nervous control, and to depend upon a direct effect of the blood-sugar level on the pancreas. This is well seen in experiments in which the pancreas has been transplanted, and therefore denervated. When glucose solution is injected into the artery supplying such a pancreas, there is a definite fall in the sugar in the circulating blood (Fig. 600). But Macleod and collaborators and G. A. Clark * have given some evidence that the liberation of insulin can also be controlled by influences reaching the pancreas by way of the vagus.

---

* Best, Hoet and Marks. *Proc. Roy. Soc.*, 1926, **100B**, 33.

† Best and Haist. *J. Phyiol.*, 1941, **100**, 142.

Nervous control does come into the picture when hypoglycæmia is present. Just as the hyperglycæmia produced by adrenaline, or in any way, can be counteracted by insulin so the hypoglycæmia brought about by an overdose of insulin can be relieved by administration of adrenaline if sufficient glycogen is present in the liver. Cannon, McIver and Bliss † showed that when insulin is given, as soon as the blood sugar falls to a certain level (about 50 mg./100 c.c.) there is excitation of the sympathetic system and a discharge of adrenaline from the suprarenals. This produces accelerated glycogenolysis in the liver, and so checks the further fall of blood sugar. We thus see that the regulation of the blood sugar involves a balance of several competing factors. Apart from the entry of carbohydrate into the body during absorption from the intestine, glycogen is constantly being produced in the liver, certainly from protein, probably also from fat, by processes of glyconeogenesis. According to the extent to which the carbohydrate requirements of the body are being met by absorption from the intestine, this glycogen is converted into glucose and liberated, rapidly in fasting, slowly during the well-fed state, into the blood passing the liver. The process of glyconeogenesis is accelerated by the diabetogenic hormone of the pituitary, partly through its action on the suprarenal, and is checked by insulin. Conversion of glycogen in the liver into glucose is accelerated by adrenaline. When the blood sugar tends to rise, more insulin is set free from the islets of the pancreas, and this not only checks glyconeogenesis

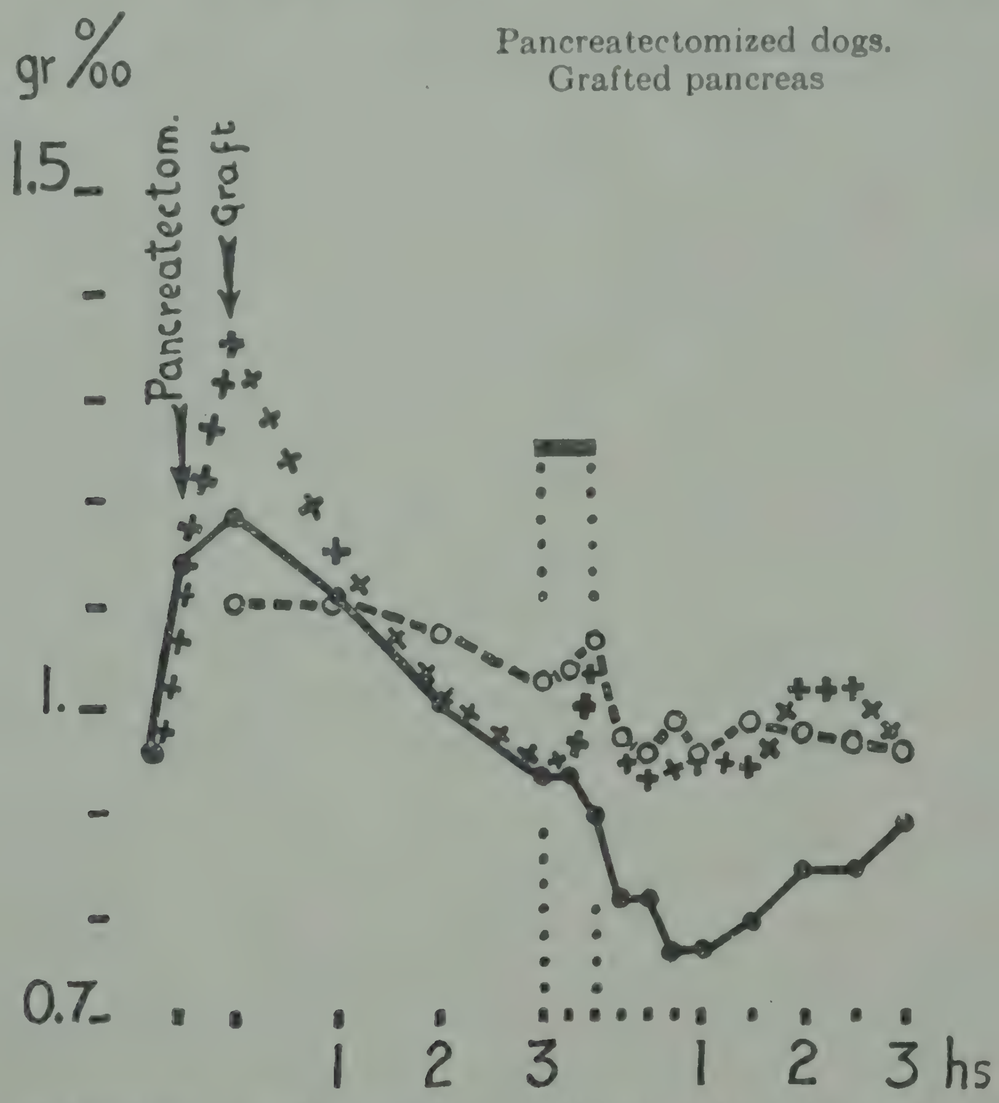


FIG. 600. Injections, for 30 minutes, at signal, into artery of transplanted pancreas in two dogs. ●▬● of 2·5 per cent. glucose. + + + of Locke's solution. Ordinates = blood sugar.

As a control 0- - - O shows effect of injection of glucose solution into jugular vein. (HOUSSAY, after FOGLIA and FERNANDES, 1935.)

---

* G. A. CLARK. *J. Physiol.*, 1926, **61**, 13, 576 ; 1927, **64**, 229 ; 1931, **73**, 297.

† CANNON, MCIVER and BLISS. *Amer. J. Physiol.*, 1924, **69**, 46.

in the liver, but also accelerates glycogen storage in the muscles and oxidation of glucose in the tissues. When the blood sugar tends to fall, adrenaline is liberated into the blood stream, and conversion of glycogen to glucose in the liver and of glycogen into lactic acid in the muscles is speeded up, until the blood sugar level is set right again.

Removal of the pancreas not only depresses the oxidation of glucose in the tissues, but also leaves the process of glyconeogenesis in the liver unchecked. Removal of the liver, or of the pituitary, or, to a lesser extent, the suprarenal, leaves the glucose-removing functions of the pancreas in charge of the situation. Proper regulation of the blood sugar is only got by appropriate collaboration between all these various factors.

Diabetes in Man.* There is good reason to suppose that diabetes in man may be due either to excessive secretion of the pituitary "diabetogenic" hormone, or to inadequate secretion of insulin by the pancreas. The latter form seems much the more common. It has been shown by Himsworth † that there are two types of diabetic, one sensitive to insulin and one less, or more slowly, responsive. The sensitive type, if given glucose and insulin, show an almost normal response to the latter ; the insensitive type, on the other hand, shows the great rise of blood sugar in response to the glucose, as though no insulin had been given (Fig. 601). It is suggested that the insulin-sensitive types, which are usually the more severe, are due to lack of insulin in the body, while the insensitive types are due to impairment of the action of insulin on the tissues, as happens in normal individuals on low carbohydrate diets, presumably because of over-production of pituitary hormones. In its severer forms, the diabetes of man resembles very closely that produced in the dog by total extirpation of the pancreas. The output of urine is largely increased, and the urine, though light in colour, is of a high specific gravity, 1,030 to 1,035, and

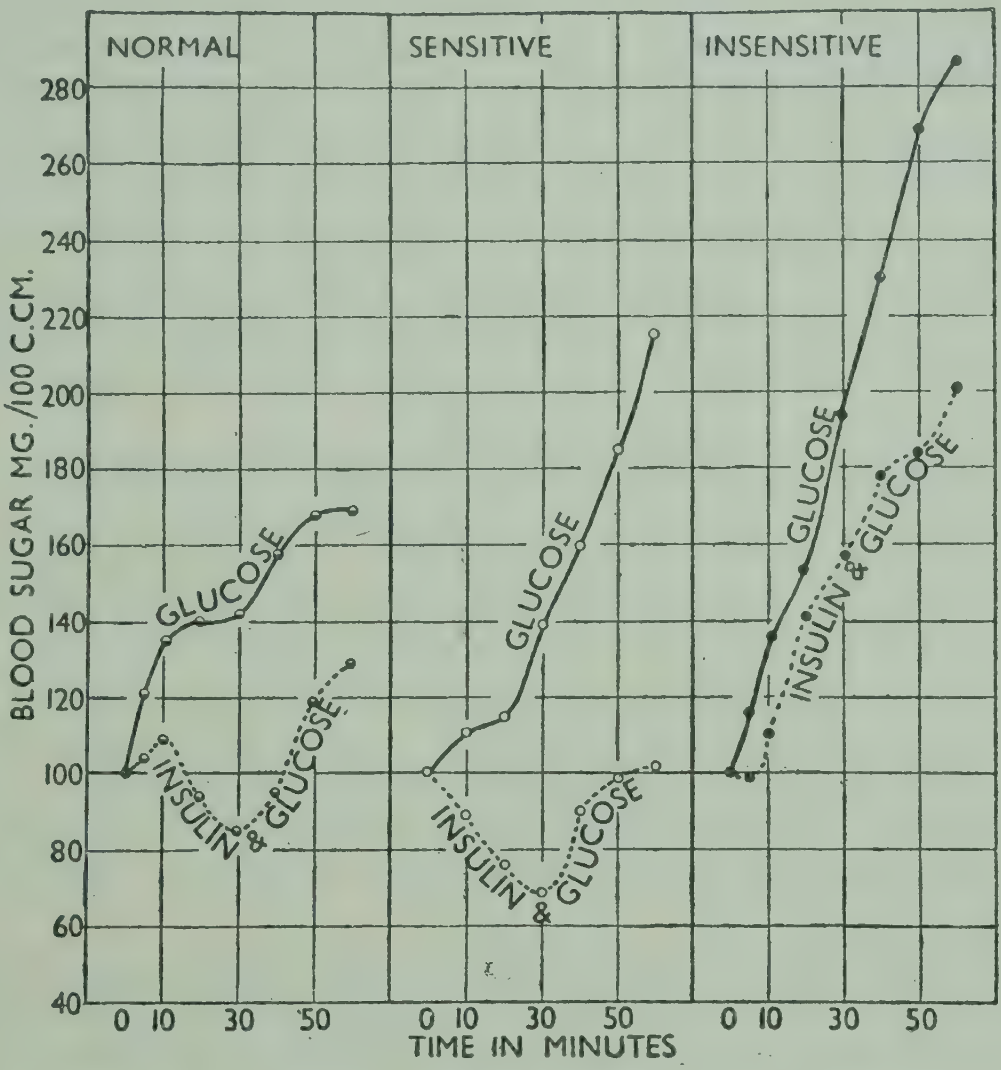


FIG. 601. Glucose tolerance and insulin-glucose curves from a normal subject, from an insulin-sensitive diabetic and from an insulin-insensitive diabetic, all on same diets. (HIMSWORTH and KERR, *Clin. Sci.*, 1939, **4**, 119.)

---

* HIMSWORTH. *Lancet*, 1949, i, 465.

† HIMSWORTH, *B.M.J.*, 1940, i, 719.

may contain from 5 to 10 per cent. of sugar. The appetite is largely increased, yet the body wastes. The patient after some months or years may die in a condition of diabetic coma. Warning of the onset of this condition is given by the rise of ammonia in the urine, and by the appearance of β-hydroxybutyric and aceto-acetic acids. The breath may smell of acetone, and this substance may also be present in the urine. The diabetic state is attended by diminished resistance of the tissues to infection. In a few cases the pancreas is found to be atrophied or diseased, but in the large majority no marked pathological change is to be observed with the naked eye in this organ. Where it has been possible to apply refined methods of microscopical examination, some change has often been found in the islets of Langerhans. We may find all grades between such cases and those in which there is still a considerable power of glucose assimilation. In order to determine the grade of the disorder, it is usual to give a glucose test meal, and to examine the urine for sugar and make a blood-sugar curve. In many cases the sugar will disappear from the urine on the administration of a diet consisting entirely of proteins and fats. When this has been effected, carbohydrates may be added in small proportions to the diet up to the limit of the assimilatory powers of the patient. Some think that administration of carbohydrate in excess of this limit is of disadvantage to the patient and hastens the progress of his disorder. When the power of assimilating carbohydrates is entirely abolished, death is inevitable unless insulin be given. This point may be determined in two ways. In the first place, a patient with no power of carbohydrate assimilation will continue to excrete sugar in the urine on a pure protein-fat diet. A very low respiratory quotient is a sign of severity of the disorder.

As in the diabetes produced by extirpation of the pancreas, the patient may often be restored to health, and kept in that condition, by the repeated administration of hypodermic doses of insulin. Increasing the carbohydrate in the diabetic diet does not necessitate a corresponding increase in the amount of insulin required, because, as with normal individuals, the carbohydrate tolerance and the insulin sensitivity are raised when the carbohydrate in the diet is increased. Care must be taken not to give too large a dose of insulin, as the hypoglycæmia thereby produced is more immediately dangerous than the hyperglycæmia. This state of hypoglycæmia, if accidentally induced by an overdose, can be removed by administration of glucose, either orally or by intravenous injection.

A drawback in insulin treatment is that, unless very frequent small injections be given, the dosage is rather sudden and irregular. Many attempts have been made to find some means of spreading out the action of a given dose of insulin so as more closely to imitate the steady production by the pancreas. The most useful of these are by the use of zinc insulinate and protamine zinc insulinate, both of which are more slowly absorbed and so produce more smooth effects. If the administration of insulin be intermitted, the patient returns at once to the diabetic condition.

For Reference

Barron (1943). The Mechanism of Carbohydrate Metabolism. *Advances in Enzymology* **3,** 149. (New York: Interscience Publ. Inc.)

Bouckaert and de Duve (1947). The Action of Insulin. *Physiol. Rev.*, **27,** 39.

Cori (1931). Mammalian Carbohydrate Metabolism. *Physiol. Rev.*, **11,** 143.

Cori (1946). Enzymatic Reactions in Carbohydrate Metabolism. *Harvey Lects.*, **41,** 253.

Deuel (1936). The Intermediary Metabolism of Fructose and Galactose. *Physiol. Rev.*, **16,** 173.

Elliott (1941). Intermediary Metabolites and Respiratory Catalysis. *Physiol. Rev.*, **21,** 267.

Haist (1944). Factors affecting the Insulin Content of the Pancreas. *Physiol. Rev.*, **24** 409.

Jensen and E. A. Evans (1934). The Chemistry of Insulin. *Physiol. Rev.*, **14,** 188.

KREBS (1943). The Intermediary Stages in the Biological Oxidation of Carbohydrate. *Advances in Enzymology*, **3,** 191. (New York : Interscience Publ. Inc.)
LAZAROW (1949). Factors Controlling the Development and Progress of Diabetes. *Physiol. Rev.*, **29,** 48.
LEWIS (1949). Diabetes and the Insulin-administration Problem. *Physiol. Rev.*, **29,** 75.
LONG and WHITE (1938). Intermediary Carbohydrate Metabolism. *Ergebn. d. Physiol.*, **40,** 164.
LUKENS (1948). Alloxan Diabetes. *Physiol. Rev.*, **28,** 304.
PETERSON (1934). The Metabolism of the Eviscerate Preparation. *Physiol. Rev.*, **14,** 586.
RAPPORT (1930). The Interconversion of the Major Foodstuffs. *Physiol. Rev.*, **10,** 349.
RUSSELL (1938). Relation of Anterior Pituitary to Carbohydrate Metabolism. *Physiol. Rev.*, **18,** 1.
SHAFFER (1923). Intermediate Metabolism of Carbohydrates. *Physiol. Rev.*, **3,** 394.
SOSKIN (1941). The Blood Sugar : its Origin, Regulation and Utilization. *Physiol. Rev.*, **21,** 140.
SOSKIN and LEVENE (1946). Carbohydrate Metabolism. (Chicago Press.)
STADIE (1942). Intermediary Metabolism in Diabetes Mellitus. (*Harvey Lects.*)
STAMMERS (1926). Recent Advances in the Study of Blood Sugar and Diabetes. *Physiol. Rev.*, **6,** 630.
SWANN (1940). The Pituitary-Adrenocortical Relationship. *Physiol. Rev.*, **20,** 493.
WOOD (1946). The Fixation of Carbon Dioxide and the Inter-relationships of the Tricarboxylic Acid Cycle. *Physiol. Rev.*, **26,** 198.
YOUNG (1948). The Mechanism of the Action of Insulin. *Science Progress*, No. 141, 13.

BOOK VII

THE REMOVAL OF WASTE MATERIAL AND TEMPERATURE CONSERVATION

EXCRETION AND TEMPERATURE REGULATION

HISTORICAL NOTE

Kidneys and Urine. Knowledge of the functions of the kidneys is inseparable from that of their structure, and of the chemical constituents of the urine. Galen, about A.D. 200, knew, and perhaps discovered, that the urine is formed by the kidneys, but it was centuries before knowledge progressed much further than that. Early in the 17th. century an attempt was made by Van Helmont to discover the nature and origin of the urinary constituents, and in 1669 the alchemist Brand isolated elementary phosphorus from urine, and was doubtless astonished by its properties; but proper appreciation of the nature of the urinary constituents had to await the development of organic chemistry, and, indeed, played a part in that development, during the 19th century. There is thus little reason for surprise at the failure of Boerhaave's attempt, early in the 18th century, to make an analysis of the urine. The first rough quantitative analysis was, in fact, made by Berzelius in 1809.

The first step towards realisation of the chemical nature of the urine was the discovery in it of hitherto unknown substances. Of these, urea, recognised by Rouelle in 1773, but fully discovered in 1797 by Cruickshank, was of critical importance. By 1822, Prévost and Dumas had shown that after nephrectomy the blood contained large amounts of urea, and this gave grounds for believing that the kidneys merely removed from the blood substances already there. Urea was of further historical importance as the first organic substance to be prepared synthetically (Wöhler, 1828): it was shown by Liebig in 1842 to arise from ingested protein, and by Meissner, about 1868, to be formed in the liver.

The excretion of hippuric acid after ingestion of benzoic acid was observed by Wöhler in 1824, and was the first discovered case of a synthesis in the animal body. But when Schmiedeberg, in 1876, showed that the synthesis was effected in the kidney, somewhat of a check was suffered by the opinion that the kidney was solely an excretory organ.

Uric acid had been isolated by Scheele and Bergmann (1776), and its mode of origin studied by Kossel (1885) and by Horbaczewski (1890). Thereafter the story may be left as recent history.

Knowledge of the details of kidney structure accumulated side by side with that of the chemistry of the urine. The uriniferous tubules, seen by Bellini in 1662, the glomeruli injected by Malpighi in 1665, and the terminal capsules observed by J. Müller in 1830 were all brought into proper relationship in 1842 by Bowman, who promulgated the first plausible theory of urine formation. This was the filtration-secretion theory, which is referred to more fully in the text of the next chapter. Two years later, however, the great mechanist Ludwig enunciated an opposing theory, the filtration-reabsorption theory. Although Goll's work, in 1854, lent colour to this theory by showing that the urine flow was closely related to the arterial pressure, many facts were left unexplained, both as regards the vascular effects, and as regards the relative concentrations of the urinary substances in blood and urine; and when Ustimowitsch, in 1869, showed that the injection of urea would re-start urine secretion previously checked by low arterial pressure, Ludwig was prepared to admit that the composition, as well as the pressure, of the blood was an important factor.

Ludwig's customary antagonist, Heidenhain, supported Bowman, and by his classical experiments with indigo-carmine (1875) considered that he had finally vindicated the tubular secretion theory, for which further evidence was provided by Nussbaum's well-known experiments on the frog kidney in 1878.

In this state of conflicting indecision the matter remained until the third decade of the present century. In the intervening period a good deal of work was done, such as that on the relation between blood pressure, kidney volume (with the oncometer devised by Roy for the purpose in 1881), blood flow, and urine flow, but none of it decisive as between the two contending theories. Cushny, in 1917, showed that, by the hypothesis of threshold substances, the filtration-reabsorption theory could be made to fit most of the facts, but it was thought by many that this exposition did not amount to demonstration.

About 1925, however, the phase of micro-experimentation was ushered in by RICHARDS' work on the frog kidney ; it became possible directly to collect and analyse the glomerular filtrate and the urine from other parts of the tubule. Work on various clearances, which began with that of urea by VAN SLYKE in 1928 and was extended to that of other substances by REHBERG (1926), SMITH, SHANNON and others, also threw much light on the whole subject. The study of the aglomerular kidneys of fishes by MARSHALL and others from 1928 also gave a new viewpoint by bringing positive proofs of normal tubular excretion. The present position is that the filtration-reabsorption theory is the one chiefly upheld for the mammalian kidney, but that in specific instances tubular secretion also occurs.

Acquaintance with the movements of the urine, and the mechanism of micturition mainly developed during the great growth of experimental physiology in the late 19th century. HALLER, in 1778, recognised that urine was retained by a reflexly maintained tone. The active period began with KOHLRAUSCH's work on micturition in 1854 and ENGELMANN's work on the movements of the ureters (1869). The innervation was mainly settled by the studies of BUDGE (1872), ZEISSL (1893), and LANGLEY and ANDERSON (1896) : the reality of the internal sphincter was confirmed by REHFISCH in 1897.

Skin and Heat Regulation. The structure of the skin was described by MALPIGHI (1687), who noted the sweat glands, more accurately described by PURKINJE in 1834, and enumerated in different areas of the body by KRAUSE in 1844.

The insensible perspiration was first studied by SANCTORIUS (1614), and afterwards by DODART, 1725, LAVOISIER and SEGUIN, 1790, and others, right up to modern times. Sweat was collected by ANSELMINO in 1844, and first analysed by SCHOTTIN in 1851. The secretion of sweat by stimulation of a mixed nerve was first noted by GOLTZ in 1875 ; it was OSTROUMOW and LUCHSINGER who showed, in 1876, that the nerve supply was of sympathetic origin, the exact path for which was mapped out by LANGLEY in 1891.

That animals were separable into those with constant and those with variable temperatures, rather than into warm and cold-blooded divisions, was pointed out by JOHN HUNTER in 1837, and again by DONDERS in 1845 : the names homoiothermic and poikilothermic respectively were introduced by BERGMANN in 1847.

The invention of the thermometer by GALILEO in 1595 soon led to its application, by SANCTORIUS, to the study of fevers. They used air thermometers, and a fanciful form of clinical air-thermometer was invented by the Grand Duke Ferdinand of Tuscany in 1654. Galileo used the first liquid (alcohol) thermometer in 1612, and in 1701 NEWTON proposed that body temperature be taken as a fixed thermometric point. A mercury maximum clinical thermometer was introduced by EHRLE in 1869. Thermo-electric couples for temperature measurements were used by ROSENTHAL in 1876, and the radiation of heat from the skin was first studied by the use of the thermopile by CHRISTIANI and KRONECKER in 1878.

A survey of body temperatures and of the topography of skin temperatures was made by J. DAVY in 1845, and it was from his studies that the mean temperature in the mouth was accepted as 98·4° F.

A series of calorimeters has been used for the study of animal heat ; the ice calorimeter of LAVOISIER and LAPLACE (1780) was followed by a water calorimeter by CRAWFORD in 1788, which was further improved by DULONG in 1842. An air calorimeter, the forerunner of many later types, was tried out by SCHARLING in 1849, and D'ARSONVAL in 1886 invented a differential air calorimeter. The modern highly accurate types are mostly of the ATWATER and BENEDICT pattern, and date from about 1905 : they are elaborations of the water-calorimeter.

It was LAVOISIER who showed, in 1780, that the heat generated in the body could be accounted for by the combustion processes which terminated in the production of CO_2 and water ; many others, notably RUBNER (1894) fully confirmed this, and showed that the heat produced was the same as that formed by combustion of the ingested food materials. The chemical regulation of heat production was first postulated by VOIT in 1878.

The ancients could in no way explain the warmth of animals, and considered it to be an innate " vital " phenomenon ; ARISTOTLE (350 B.C.) believed it to emanate from the heart, and that breathing cooled the body. There was not much advance in opinion until late in the 18th century. Thus, VAN HELMONT (1648) still attributed it to the " vital spirits " of the blood, DESCARTES and WILLIS (1670) to a fermentation, BOERHAAVE (1708) and STEPHEN HALES (1733) to friction of the blood, and HALLER (1760) held some similar opinion.

JOHN MAYOW, certainly, got rather near the modern idea in 1674, with his view that the heat was due to the combination of the " nitro-aerial spirit " (= oxygen) with combustible " particles ". But it was LAVOISIER's demonstration of the real nature of combustion, in 1777, that brought the subject into the ambit of modern views. He was wrong, however, in supposing that the lungs were the seat of the oxidative processes, and it was not until it was shown that the blood of the right side of the heart is warmer than that of the left (MALGAIGNE, 1832, BERGER, 1833, and many since) that this error was brought home. The demonstration that in deep-seated organs the venous blood is warmer than the arterial (LIEBIG, 1853, CL. BERNARD, 1876, etc.), together with the work of PFLÜGER and others showing that the tissues were the seat of the oxidative processes, fully confirmed this viewpoint. The notion that, nevertheless, the lungs were capable, when the tissues had failed, of completing oxidative processes, was again put forward in 1897 by Bohr and Henriques, but was speedily repudiated.

Boerhaave thought that lung-breathing animals could not survive in surroundings at temperatures above their own : this was opposed by Haller, on the evidence of tropical explorers, and Ellis, in 1758, made the significant observation that the human skin had a temperature of only 97° F. when the environmental temperature was 105° F. Franklin explained this as due to the evaporation of the sweat. The relative effects of dry and moist air were shown by Delaroche and Berger in 1806, and in 1875 Blagdon and others performed the crucial experiment of supporting unscathed a temperature of 127° C. in dry air for several minutes.

That the body temperature was influenced by the central nervous system became evident from the experiments of Benjamin Brodie in 1811, and was confirmed by the subsequent work of Williams (1835), Claude Bernard (1876), and many others. Damage to the spinal cord or other parts of the C.N.S. often caused a fall, but sometimes a rise, of body temperature : hyperthermia was especially noted by Tscheschicin, in 1866 (damage to pons), and by Aronsohn and Sachs (1885) and Hale White (1890), who first showed that puncture in the region of the corpus striatum produced a prolonged rise of body temperature.

CHAPTER XLVII

RENAL EXCRETION

1. THE COMPOSITION AND CHARACTERS OF THE URINE

WATER is got rid of by the lungs, skin, and kidneys. The residues of the bile and digestive secretions are eliminated with the fæces. With these exceptions, practically all the waste products, except CO_2, resulting from metabolism are excreted in the urine. We have thus in this fluid the last chapter in the metabolic history of a large number of the constituents of the diet. Since, moreover, the kidneys may excrete almost any substance which circulates through their blood vessels, many of the intermediate metabolites may be found in minute quantities in the urine. Under pathological conditions, these metabolites may appear in the urine in larger amounts and serve then as an index to some interference with metabolism.

The composition of the urine must, therefore, vary according to the activity of the body, the quantity and nature of the food taken, and the relative amount of water escaping by other channels.

The urine of man is a clear yellow fluid, which froths when shaken. On standing, a cloud of mucus is deposited, consisting of a very small amount of protein derived from the epithelial lining of the bladder and urinary passages. In concentrated urine a deposit of *urates* occurs on cooling, and dissolves when the urine is warmed. When normal urine is turbid as it is passed, the turbidity generally consists of calcium phosphate and is not cleared up by heating, but dissolves in dilute acid.

The *colour* of the urine varies with its concentration. After severe sweating the urine is of high colour and after copious draughts of liquid may be very pale.

Ordinary urine has an aromatic *odour*, but this varies largely with the character of the food.

The *specific gravity* of the urine is proportional to its concentration. Normally it is 1,016 to 1,020, though it may rise as high as 1,040 or sink as low as 1,002. The specific gravity of the urine when fluid intake is suspended is an important index of the functional condition of the kidney. If fluid intake is stopped for twenty-four hours, the urine secreted during the second half of the period should have a sp. gr. not below 1,026 : in nephritis it may be as low as 1,010.

Its osmotic pressure is usually much above that of the blood plasma, Thus, the Δ of urine normally varies between 0·87 and 2·71° C. (Δ of blood = 0·56°), but after copious draughts of water the Δ may be as small as 0·25° C.

The *reaction* of urine is generally acid. This is due to the fact that neutral constituents of the food may give rise to acid end-products of metabolism. The sulphur of proteins is converted into sulphuric acid and the phosphorus of nucleic acid and lecithin into phosphoric acid. There is thus a predominance of H-ion over OH-ion in ordinary urine. The *p*H of the urine can easily be determined by indicators, is on the average 5·3, and varies from 8·0 to 4·8.* This, however, applies only to man and to carnivora. In the food of herbivora there is a predominance of salts of organic acids. The acids, *e.g.* tartaric, malic, and citric acids, undergo oxidation in the body, so that their corresponding cations have to be excreted ; hence the H-ion concentration of the urine is reduced ; the urine of herbivora therefore is alkaline in reaction, and turbid. If a herbivorous animal be starved, so that it has to live on its own tissues, its urine becomes clear and acid. The urine of man can similarly be made alkaline by the ingestion of large quantities of vegetables, by the administration of alkalies (*e.g.* $NaHCO_3$) by the mouth, or by prolonged forced breathing.* Under such circumstances, the urine, as passed, is turbid from the presence of calcium phosphate.

---

* HENDERSON and PALMER. *J. Biol. Chem.*, 1912, **13**, 394.

In determining the titratable acid of the urine, it is usual to adhere to one indicator, *e.g.* phenolphthalein, and to give the acidity in terms of decinormal acid, naming the indicator used. The total acid set free in the body per day is equal to about 700 c.c. of N/10 acid, but some is neutralised and excreted as ammonium salts (= c̄ 400 c.c. N/10 acid daily). The rest, the urinary titratable acid, amounts to about 300 c.c. $\frac{N}{10}$ acid daily.

Owing to the buffering action of phosphates, the *titratable acidity* of urine is always much greater than the H-ion concentration would suggest. By titrating, first with methyl red and an acid, and then with phenolphthalein and an alkali, we obtain a useful measure of the buffers of the urine. By the first titration we convert the dibasic phosphates to the monobasic salt (NaH_2PO_4), and by the second we convert the monobasic to the dibasic phosphate (Na_2HPO_4). Hence the ratio of the first to the second titration gives us the ratio $\frac{Na_2HPO_4}{NaH_2PO_4}$.† Other buffer systems are also present.

For Reference

BRUNTON (1933). The Acid Output of the Kidney and the So-called Alkaline Tide. *Physiol. Rev.*, **13**, 372.

THE AVERAGE COMPOSITION OF THE URINE. The following may be taken as a fair daily average for an adult man on ordinary mixed diet, in Europe :

Total amount of urine = 1,500 c.c.

This contains about 60 g. of solids, of which 25 g. are inorganic and 35 g. organic. The chief constituents are distributed roughly as follows :

INORGANIC CONSTITUENTS	
Sodium chloride	15·0 g.
Potassium	3·3 ,,
Sulphuric acid	2·5 ,,
Phosphoric acid	2·5 ,,
Ammonia	0·7 ,,
Magnesium	0·1 ,,
Calcium	0·3 ,,
Other substances	0·2 ,,
Iron	0·005 ,,

ORGANIC CONSTITUENTS	
Urea	30·0 g.
Creatinine	1·0 ,,
Uric acid	0·7 ,,
Hippuric acid	0·7 ,,
Indican	0·01 ,,
Acetone bodies	0·04 ,,

The quantity of urine will naturally vary according to the habit of the individual with regard to the intake of fluids and with his occupation. Thus, after copious sweating, the total amount may fall to 400 c.c. in the course of the day. If large draughts of liquid be taken, it may rise to 3,000 c.c. or more. There are also diurnal variations in the amount secreted, probably depending largely on the circulation through the kidneys. The secretion is at a minimum during sleep, and especially between 2 and 4 o'clock in the morning. It is at its maximum during the first hours after rising, and generally increases after each meal. Muscular exercise may also give an initial increase, owing to the greater vigour of the circulation associated with exercise. If the exercise is severe enough to cause sweating, or is carried to fatigue, there may be a consequent diminution in the amount of urine secreted.

THE INORGANIC CONSTITUENTS OF THE URINE

(*a*) **Anions.** The *chlorides* of the urine are derived from the chlorides of the food. The output of chlorides, which normally varies from 6 to 10 g. Cl. per day, will therefore vary with the intake. On Cl.-free diet the chlorides may almost disappear from the urine, although the blood contains practically the normal amount of chlorides, showing that the body retains essential chloride as long as possible. Chlorides may also temporarily disappear from the urine under various pathological conditions, *e.g.* acute pneumonia.

Sulphates. The sulphates of the urine are derived almost entirely from the oxidation of the sulphur of protein. Hence, as the nitrogen of the urine goes up, so the sulphates will increase. On an average diet, the ratio of urinary nitrogen to SO_4 is about 5 : 1 ; though this ratio will alter with the nature of the protein taken as food. The daily output of sulphuric acid varies between 1·5 and 3 g. SO_4. The greater part is present as sulphates of the alkaline metals. A certain proportion, about 10 per cent., occurs in the form of ethereal sulphates, chiefly indoxyl sulphate. It is formed in the liver by the union of phenolic compounds with the sulphate ion.* A small proportion of the sulphur in the urine is as so-called neutral sulphur. This probably includes a number of different bodies, among which thiocyanates, mercaptans and cystine are the best known.

---

* DAVIES, HALDANE and KENNAWAY. *J. Physiol.*, 1921, **54**, 32.

† LEATHES, *B. M. J.*, 1919, p. 165 ; *Lancet*, 1920, ii., 933.

Phosphates. The phosphates of the urine are derived, partly from the phosphates of the food, partly from the oxidation of the organic phosphorus-containing constituents of the food and of the tissues, *e.g.* nuclein, lecithin, etc. If the food contains much calcium and magnesium, the amount of phosphates excreted in the urine diminishes, since they are excreted with the fæces as calcium and magnesium phosphates. According to the diet therefore, phosphoric acid may be excreted either by the intestine or by the kidneys. The amount of phosphates, reckoned as P_2O_5, excreted in the course of the day may vary between 1 and 5 g. In the urine, the phosphates exist as a mixture of the mono- and di-basic phosphates (Na, K, NH_4, etc.), the relative amounts of the di- and mono- being dependent upon the hydrogen ion concentration of the urine. If the urine is neutral or alkaline there is often a deposit of earthy phosphates. Whether this deposit is present or not, depends on the reaction of the urine. Thus the mono-magnesium phosphate $MgH_4(PO_4)_2$ and the mono-calcium phosphate $CaH_4(PO_4)_2$ are both fairly soluble in water, and would be present in acid urine. With lowered acidity of the urine, the proportion of the two bases present in these forms is diminished. The di-magnesium and di-calcium phosphates are only slightly soluble in water, and would be deposited. One may indeed, in slightly acid urine, find the di-calcium phosphate occasionally present as a crystalline deposit (" stellar phosphate "). On heating the urine, the di-calcium phosphate breaks up into a mono-calcium phosphate and a tri-calcium phosphate, while the acidity of the urine is increased by the solution of the mono-calcium phosphate. Alkaline urine will always present a precipitate of tri-calcium phosphate, $Ca_3(PO_4)_2$. When normal urine is allowed to putrefy, the urea is converted into ammonia, and the urine becomes alkaline. Under such conditions we may often find a crystalline precipitate of ammonium magnesium phosphate, NH_4MgPO_4, the so-called " triple phosphate."

(*b*) **Cations.** The principal ones are potassium, sodium, ammonium, magnesium and calcium.

The amount of *potassium* excreted per day varies between 2 and 3 g., according to the food. With a large meat diet, and in fasting, the output is increased.

The amount of *sodium* excreted in the twenty-four hours is on the average about 5 g., but fluctuates largely with the diet.

Calcium and *magnesium* are invariably present in urine, but in much smaller quantities than the alkaline metals. The average amount of these two bases in the twenty-four hours varies in each case between 0·1 and 0·2 g.

Normal human urine always contains a small amount of *ammonia*, on an average between 0·6 and 0·8 g. in the twenty-four hours. As we have already seen, in dealing with the origin of urea, the quantity of ammonia in the urine is an index to the excess of acids over bases which have to be excreted. Nash and Benedict † showed that the ammonia of the urine is almost entirely formed in the kidney itself, and does not represent an excretion from the blood. It appears to be formed mainly from glutamine, the amide of glutamic acid, which is present in small amount in the blood (except that of the renal vein).*

---

* HELE. *Bioch. J.*, 1931, **25,** 1736.

† NASH and BENEDICT. *J. Biol. Chem.*, 1921, **48,** 463 ; BENEDICT and NASH, *Ibid.*, 1929, **82,** 673.

Traces of *iron* may be excreted by the urine, as a rule less than 5 mg. per day.

ORGANIC CONSTITUENTS OF THE URINE

Of the innumerable organic constituents, many contain nitrogen ; in man on an ordinary mixed diet, most of this is distributed as follows :

Urea	85–90 per cent.
Ammonia	2–4 ,,
Creatinine	3 ,,
Uric acid	1–3 ,,

About 6 per cent. of the urinary nitrogen is in the form of other substances, such as hippuric acid, pigments, etc. All the organic constituents of the urine (except hippuric acid) are obtained direct from the blood ; hence they represent real excretory products. Their origin in the body has already been discussed in Chapter XLVI.

Urea, $CO\!<\!\begin{smallmatrix}NH_2\\NH_2\end{smallmatrix}$. From 20 to 35 g. per day are usually excreted. Urea may easily be prepared from urine by evaporating to dryness with animal charcoal and extracting the dry residue with acetone. On evaporation of the filtered extract, urea crystallises out.

Urea is hydrolysed by certain micro-organisms, *e.g.* the *micrococcus ureæ*, which is responsible for the ammoniacal change which occurs in urine when exposed to the air : $CON_2H_4 + H_2O = CO_2 + 2NH_3$.

This action is due to an enzyme, *urease*, which is also present in certain plant tissues, *e.g.* the seeds of the soya bean. When urea, in slightly acid solution, is incubated with ground soya bean, the urea is all converted into ammonium salts. This forms the basis of a method for the estimation of urea.†

Urea is the most important index to the protein metabolism of the body. It may be regarded as partly exogenous, partly endogenous. The greater part is exogenous, and it is formed in the liver (Chapter XLVI.).

Creatinine. Creatinine usually occurs in quantities of 1·0 to 1·5 g. in the twenty-four hours. It can be prepared from urine by precipitation as picrate, and conversion into the zinc chloride compound, which is then decomposed by ammonia.‡

Uric Acid. 2-6-8-trioxypurine. The amount of uric acid per day varies, according to the diet, from 0·1 to 2·0 g. Uric acid can be prepared from urine by precipitation, as ammonium urate, with ammonium chloride and ammonia.

It is only slightly soluble in water, but soluble in concentrated acids and in alkalies. At physiological *p*H values, uric acid acts as a weak monobasic acid so that it exists as an equilibrium mixture of uric acid and monobasic urate. In urine at the usual *p*H, and in the brick-red amorphous precipitate which occurs in concentrated urine on cooling, the greater part of the uric acid consists of such a mixture. The more acid the urine, the greater the proportion of free uric acid.

Small traces (16 to 60 mg. per day) of *purine bases*, xanthine, hypoxanthine and adenine, also occur in urine. When tea and coffee are taken the methyl purines may occur, namely, caffeine, theobromine and their derivatives.

Allantoin is only present in traces in human urine (5 to 15 mg. per day).

Amino-Acids. Amino-acids are always present in traces in the urine, and contribute per day from 0·1 to 1·0 g. of nitrogen.

Hippuric Acid or benzoyl glycine, $C_6H_5CO{\cdot}NH{\cdot}CH_2COOH$, is a frequent, though not a constant, constituent of human urine. It is derived from benzoic acid, or from substances which, on oxidation, give rise to it.

---

* Van Slyke, *et al.* *J. Biol. Chem.*, 1943, **150,** 481.

† Van Slyke and Cullen. *J. Biol. Chem.*, 1914, **19,** 141 ; Youngburg. *J. Biol. Chem.*, 1921, **45,** 391.

‡ Benedict. *J. Biol. Chem.*, 1914, **18,** 182.

In the dog this coupling can be effected in the kidneys, but in most other animals, including man, it is probable that the synthesis occurs mainly in the liver. In fact, in diseases of the liver, administration of benzoic acid is followed by a slower excretion of hippuric acid than is seen in normal subjects, and this has been made the basis of a liver-function test.*

The amount of hippuric acid excreted in the day may vary between 0·1 and 1 g. After a diet rich in fruit or vegetables its amount may rise to 2 g. It is present in considerable quantities in horses' urine. It forms milk-white crystals, which are only slightly soluble in cold water, but easily soluble in alcohol and in ethyl acetate.

Other Aromatic Substances. The chief of these (about 5 to 20 mg. per day) is the so-called "urinary indican" or potassium-indoxyl-sulphate. This is derived from the indol produced in the intestines from tryptophan by putrefaction (p. 930). It is converted into indigo by the action of mild oxidising agents.

The Urinary Pigments. Normal urine gives no definite absorption bands. It owes its colour to the presence of a yellow pigment, *urochrome*. In order to separate urochrome, urine is acidified slightly with acetic acid and extracted with $\frac{1}{10}$ its volume of *n*-butyl alcohol. After washing and concentrating the alcoholic solution, it is again washed with chloroform, benzene and ether; then after further concentration the solid urochrome crystallises out as a brown powder.† The pigment is precipitated by various salts of heavy metals, and is decolourised by reduction with zinc and HCl. About 73 mg. of urochrome is excreted per day by an average adult, the amount varying with the basal metabolism. Urochrome, on treatment with aldehyde, yields a pigment closely similar to urobilin. On the other hand, urobilin, evaporated with wet ether, is converted into urochrome. Urochrome is therefore probably derived from urobilin, possibly by union with a polypeptide.

Urobilin is only present in traces in normal urine, and then chiefly in the form of a chromogen, *urobilinogen*, from which it must be set free by acidification. In certain pathological conditions, especially in cirrhosis of the liver, urobilin may occur in the urine in considerable quantities.

The urobilin of urine, which has now been obtained in pure crystalline form, is closely related to stercobilin, the colouring matter of the fæces.‡ It is formed in the intestines by the reducing action of micro-organisms on bilirubin, and urobilinogen is a result of further reduction.

Urobilin may be detected in the urine by Schlesinger's test. The urine is acidified with acetic acid and extracted with amyl alcohol. The amyl alcohol layer is removed and a few drops of alcoholic zinc chloride added, when a green fluorescence appears.

Urobilinogen is detected by Ehrlich's test—addition of a solution of *p*-dimethyl-aminoazobenzaldehyde in 50 per cent. HCl gives a red colour.

Uroerythrin gives the pink colour to urate sediments. Its chemical nature is not known. It is distinguished by the fact that on addition of caustic soda the pink colour is changed to green. On suspending the red-coloured precipitate of urates in hot water and extracting with amyl alcohol, a pink solution is obtained which shows two absorption bands in the green part of the spectrum.

Porphyrins of series I (coproporphyrin I, and uroporphyrin I) are present only in very small amounts in normal urine, but under certain conditions, especially after poisoning with sulphonal, may occur in such large quantities as to give the urine a deep purple colour. Under these circumstances they are found in the form of alkaline porphyrins and give the characteristic absorption bands of the latter.

Non-Nitrogenous Organic Constituents. Among these is *oxalic acid*, of which about 30 mg. per day is excreted, though more is present if oxalates are taken by mouth; if soluble calcium salts, or milk, are taken by mouth with oxalate, or with foods (*e.g.* rhubarb) containing it, the oxalate is precipitated and not absorbed, so that the oxalate excretion by the urine is not increased.* There is also *lactic acid*, the amount of which, usually 75 to 200 mg. per day, varies with the amount of muscular exercise recently taken. Glucuronides, various derivatives of sex hormones, vitamins, etc., and traces of aceto-acetic and β-hydroxybutyric acids are also present.

---

* Quick. *Arch. Int. Med.*, 1936, **57**, 544.

† Drabkin. *J. Biol. Chem.*, 1930, **88**, 433, 443.

‡ Watson. *Ztschr. f. physiol. Chem.*, 1935, **233**, 39.

ABNORMAL CONSTITUENTS OF THE URINE †

A large number of substances occur in minute traces in the urine. Most of the so-called pathological constituents may be detected in traces in normal urine. It is only when they occur in easily detectable amounts that their presence becomes of significance.

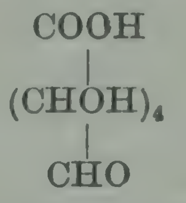
Coagulable Protein. Normal urine contains only minute traces of dissolved protein (av. about 4 mg./100 ml.). When foreign proteins of a molecular weight lower than about 68,000 are present in the blood, these are excreted in the urine; proteins of higher molecular weight, though foreign to the blood stream, are not passed out by the normal kidney.‡ If the kidneys are damaged by disease, by interference with their blood supply or by circulating poisons, the glomerular epithelium permits the passage of a certain amount of the serum albumin and serum globulin of the plasma, and on histological examination the coagulated protein may be seen in Bowman's capsule. The presence of coagulable protein (generally spoken of as albumin) in the urine is associated with Bright's disease. A small trace will often be found in the urine which is passed shortly after taking muscular exercise, but this has no pathognomonic significance.§

Hæmoglobin or its derivatives is present in the urine when hæmolysis occurs in the blood stream or urinary passages (*v.* p. 998).

Sugar. In diabetes, glucose appears in large quantities in the urine. In renal glycosuria glucose appears in the urine, but the blood sugar is rather lower than normal. Lactose may occur in the urine of nursing women.

In rare circumstances, fructose or a pentose may be found in the urine.

Glucuronic Acid. Glucuronic acid is an oxidation product of glucose, having the formula:



$$\begin{array}{c} COOH \\ | \\ (CHOH)_4 \\ | \\ CHO \end{array}$$

It reduces Fehling's solution and is lævorotatory.

About 150 mg. are excreted daily in normal conditions. Most of it is conjugated with other substances to form glucuronides, and this conjugation, effected in the liver, is one means by which poisonous substances when administered or when formed in the body are detoxicated. Camphor, chloral, and to some extent phenol, indol and skatol, and certain steroids related to the sex hormones, *e.g.* pregnandiol, are excreted in the form of glucuronides.

Acetone Bodies. These are *acetoacetic acid*, *β-hydroxybutyric acid* and *acetone*; they are present in traces (20 to 50 mg. per day) in normal urine, and in large amount in untreated severe diabetes, and may also occur in any condition of carbohydrate starvation, relative or absolute (*v.* p. 945).

Cystine. This is found in the urine of certain individuals.‖ The condition represents an inborn error of metabolism, and persists throughout life. In such cases, the cystine may give rise to urinary deposits or even to calculi.

Homogentisic Acid. This is dihydroxyphenyl acetic acid (*v.* p. 929).

It occurs in the urine of individuals affected with *alcaptonuria*.

Bile pigments are present in the urine in jaundice, owing to the presence of bilirubin in greatly increased amount in the blood plasma.

Bile Salts are also often detectable in the urine in obstructive jaundice, *e.g.* in jaundice due to gall-stones.

Urinary Deposits. In addition to formed elements, such as blood corpuscles, bacteria, or pus cells, which may occur in abnormal urine, the following deposits may be found:—

(*a*) *In Acid Urine.* (1) Amorphous urates occur generally as a brick-red amorphous deposit, thrown down as a concentrated urine cools. It is redissolved on warming the urine.

(2) Uric acid. Whetstone, dumb-bell, or sheaf-like aggregations of crystals, generally deeply pigmented so as to resemble cayenne pepper (Fig. 605).

(3) Calcium oxalate (Fig. 604). Colourless, transparent, highly refractive octahedral crystals.

---

* BARRETT. *Lancet*, 1942, **143**, 574.

† For further details *v.* HARRISON. Chemical Methods in Clinical Medicine. 1947. (Churchill.)

‡ KERRIDGE and BAYLISS. *Lancet*, 1932, p. 785; KERRIDGE, RUSSELL and BAYLISS. *J. Physiol.*, 1933, **77**, 386.

§ MACLEAN. Renal Disease. (London.) 1927.

‖ ROSENFELD. *Ergebn. d. Physiol.*, 1920, **18**, 118.

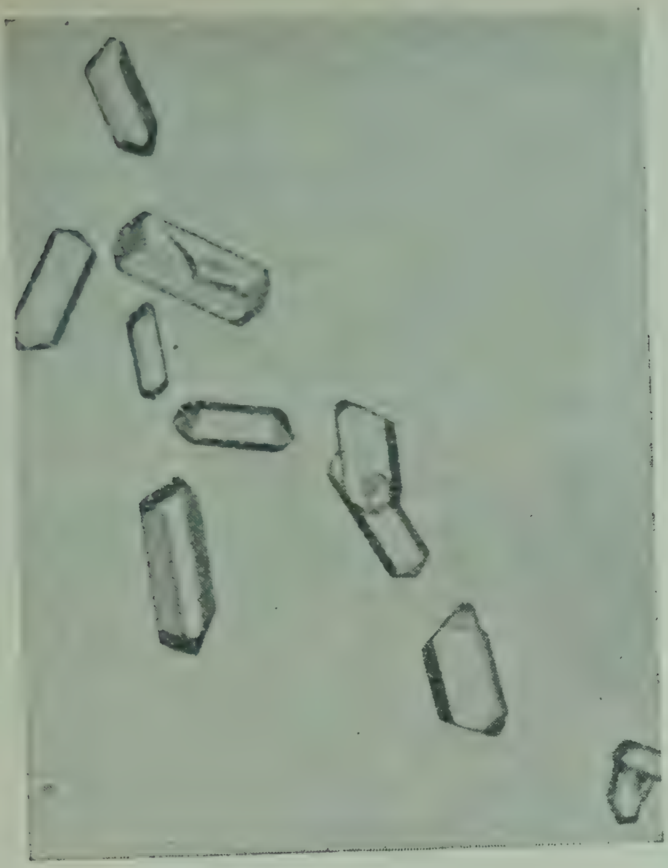


FIG. 602. Triple phosphate (knife-rest form). × 145.

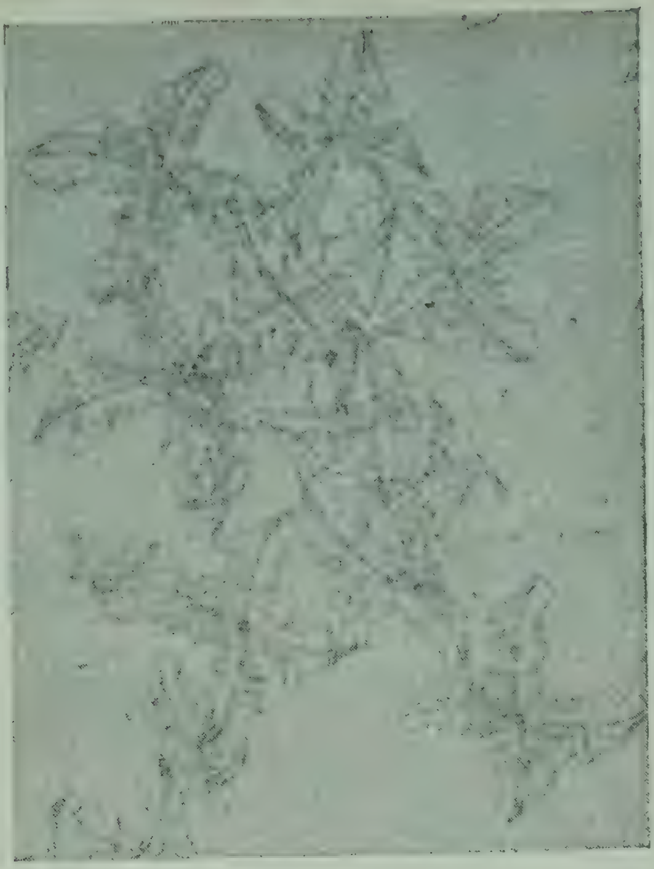


FIG. 603. Triple phosphate (feathery form). × 145.

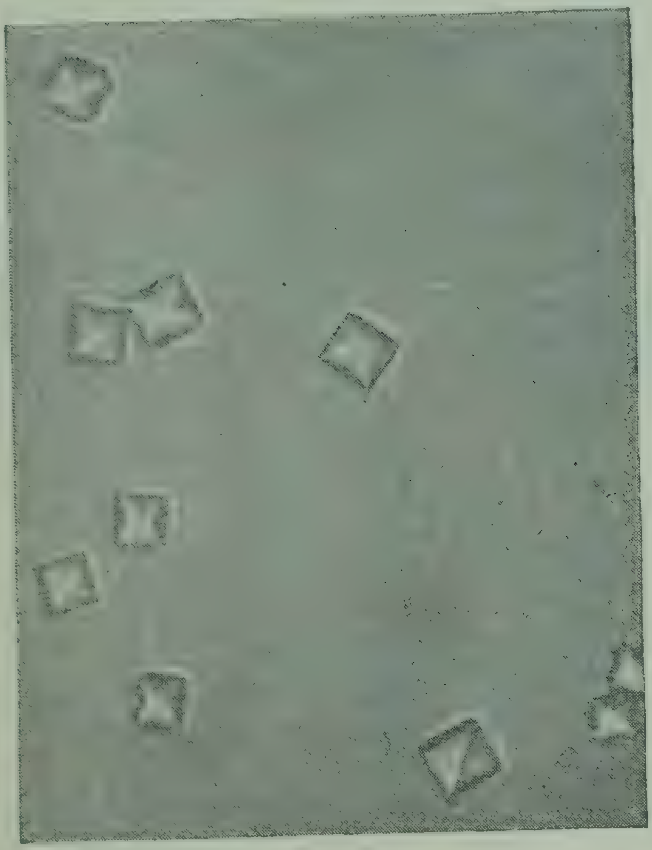


FIG. 604. Calcium oxalate (envelopes). × 640.

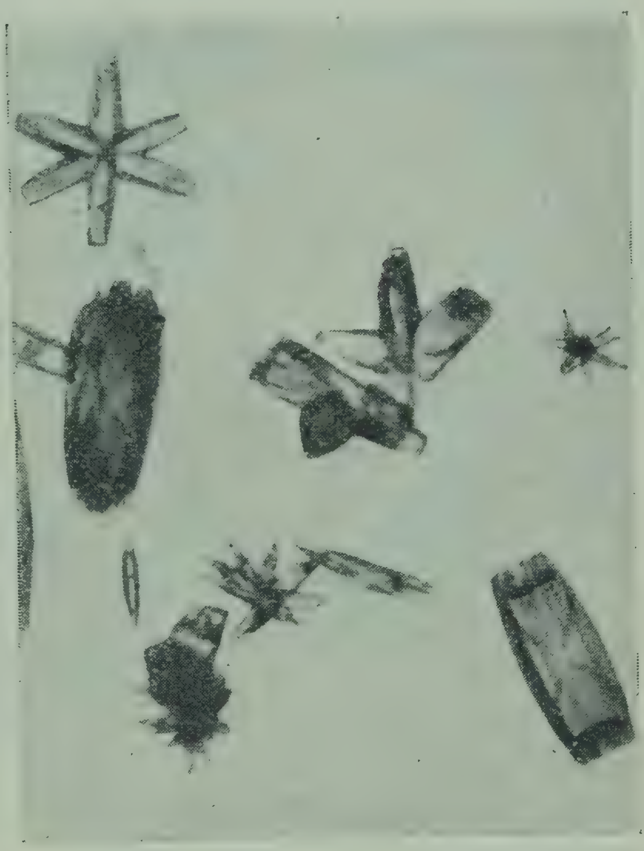


FIG. 605. Uric acid (barrels, stars, etc.). × 66.

FIGS. 602–605. Urinary deposits.

From THORPE'S "Biochemistry for Medical Students." (Churchill.)

(4) Ammonium magnesium phosphates (in faintly acid or neutral urine). The crystals have been compared to knife-rests or coffin-lids (Figs. 602 and 603).

(5) Calcium hydrogen phosphate, $CaHPO_4$. They form prismatic crystals often arranged in rosettes.

(6) Tyrosine and cystine under very rare circumstances.

(*b*) *In Alkaline Urine.* (1) The commonest precipitate consists of earthy phosphates, amorphous, easily soluble in dilute acetic acid.

(2) Ammonium magnesium phosphate, or triple phosphate (Fig. 603), is common in urine which has undergone ammoniacal fermentation.

2. THE SECRETION OF URINE

With the exception of ammonia, and perhaps hippuric acid, all the constituents of the urine are formed in parts of the body other than the kidneys, which merely excrete them from the blood. Hence extirpation of both kidneys leads to an accumulation of these specific urinary constituents in the blood and tissues. Considered from a broad standpoint, the function of this organ is the preservation of the normal composition of the circulating blood. Whenever the latter contains an abnormal non-colloidal constituent, or normal constituents in abnormal quantities, the kidney excretes the substance in question until the normal composition of the blood is restored. The problem is to determine the conditions which influence the quantity and quality of the urine secreted by the kidneys, and to ascribe to each structural element its proper share in the total work of the kidney.

Structure of the Kidney. Our views as to function are intimately dependent on our knowledge of the structure of the kidney. This organ is a branched tubular gland consisting, in man, of ten to fifteen nearly equal divisions, known as the Malpighian pyramids. In certain animals, such as the rabbit and rat, only one pyramid is present. It is divided into an outer portion, or cortex, an inner portion, the medulla, and between these the " boundary zone," containing the larger branches of the renal blood vessels (Fig. 606). From the outer boundary of the Malpighian pyramids of the medulla, a number of processes, the medullary rays, pass out into the cortex towards the surface of the kidney.

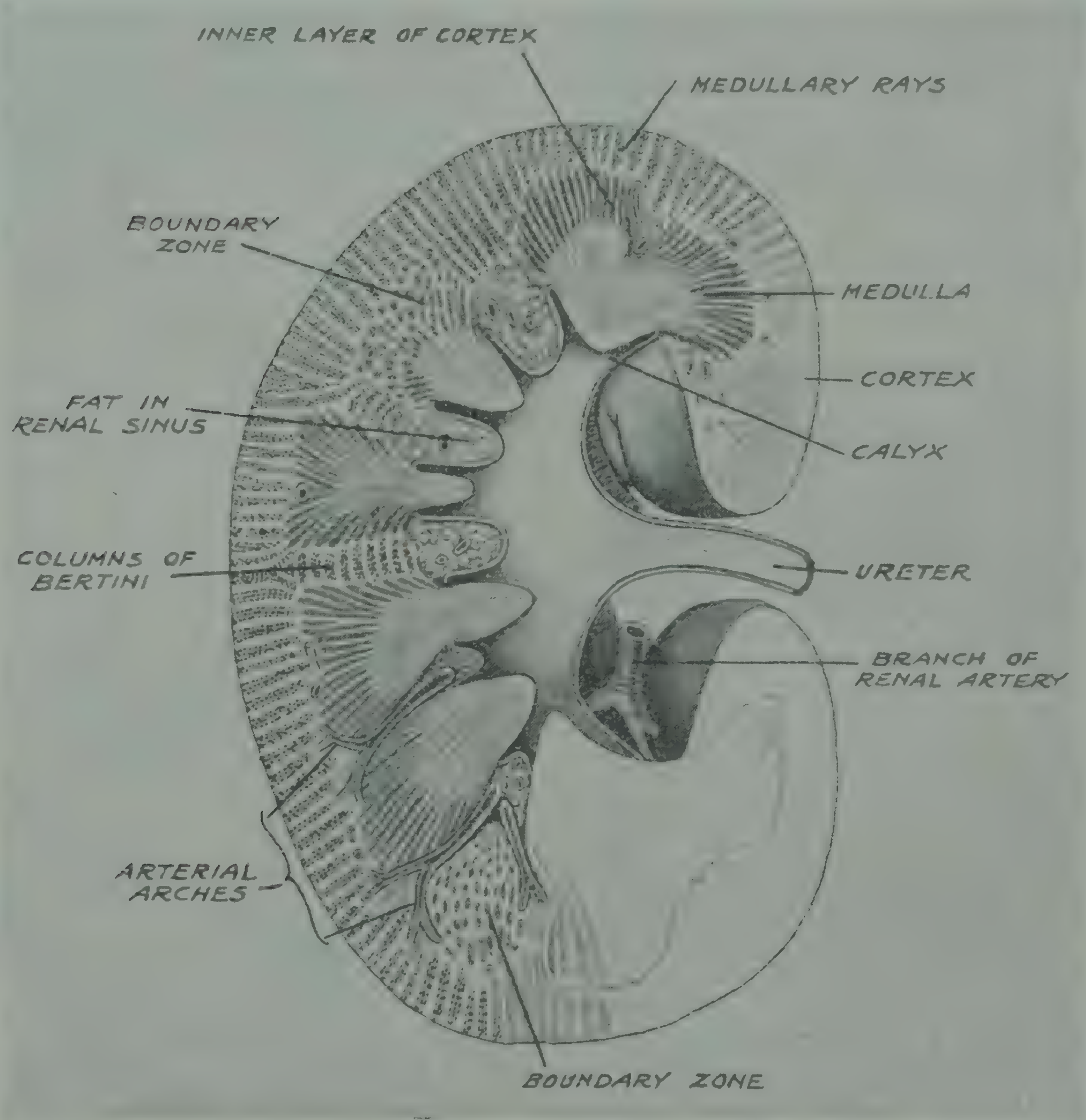


FIG. 606. Longitudinal Section of the Kidney, opening up the Kidney Sinus. The pelvis of the ureter and some of its calyces have been laid open as they lie within the sinus. (RAUBER-KOPSCH.)

The kidney is made up of separate units, or *nephrons*, all of similar structure and function. They are intricate tubules embedded in scanty connective tissue and richly supplied with blood vessels (Fig. 608). Each nephron begins by a blind dilated extremity in the cortex, known as *Bowman's capsule*, which surrounds a bunch of capillary blood vessels, the *glomerulus*, the two together forming the *Malpighian body*. In most mammals there are roughly 2 million glomeruli in the two kidneys for each sq. m. of body surface.* This means that the number is correlated with the general metabolism of the body. In man there is a greater weight of kidney, but fewer glomeruli, per sq. m. of body surface, and the blood flow to each nephron is greater than in rabbit or dog.

---

* KUNKEL. *Bull. Johns Hopk. Hosp.*, 1930, **47,** 285.

From Bowman's capsule a short *neck* (or 1st segment) leads into a *proximal convoluted tubule* (or 2nd segment), and this into a U-shaped portion which passes down into the medulla, and consists of straight *descending* (3rd segment) *and ascending limbs and the loop of Henle* (Fig. 608). The ascending limb passes into a *distal convoluted tubule* (which, together with the ascending limb of Henle, is called the 4th segment), and this by a *junctional tubule* (5th segment) joins with a number of others to form a straight *collecting tubule.* Several of these unite to form the papillary ducts, which open on the surface of the papilla in the calyx. The whole tubule consists of epithelium lying on a basement membrane; the epithelium varies in structure in different parts of the tubule. The glomerular capillaries are covered with a very thin layer of endothelial cells, and a similar layer forms the lining of Bowman's capsule. The convoluted tubules contain cells which are cylindrical in cross-section and have long been distinguished as "rodded epithelium" (Fig. 607) on account of the ease with which a radial disposition of rods or granules is demonstrated in their protoplasm. Moreover, the free margin of the cells, instead of being irregular, consists of a well-marked striated border, formed of a number of very fine processes closely set together and springing from a row of granules in the peripheral part of the cell (Fig. 607). The processes, which make up the striated border, have not been observed to present ciliary movement, and are probably comparable with the similar structures found clothing the free border of the epithelium of the intestinal villus. The descending limb of Henle's loop is narrow, and possesses flattened epithelial cells, while the ascending limb presents an epithelium similar to that of the convoluted tubules, but with less marked striation. The junctional and collecting tubules are lined with cubical or columnar cells with a clear protoplasm.

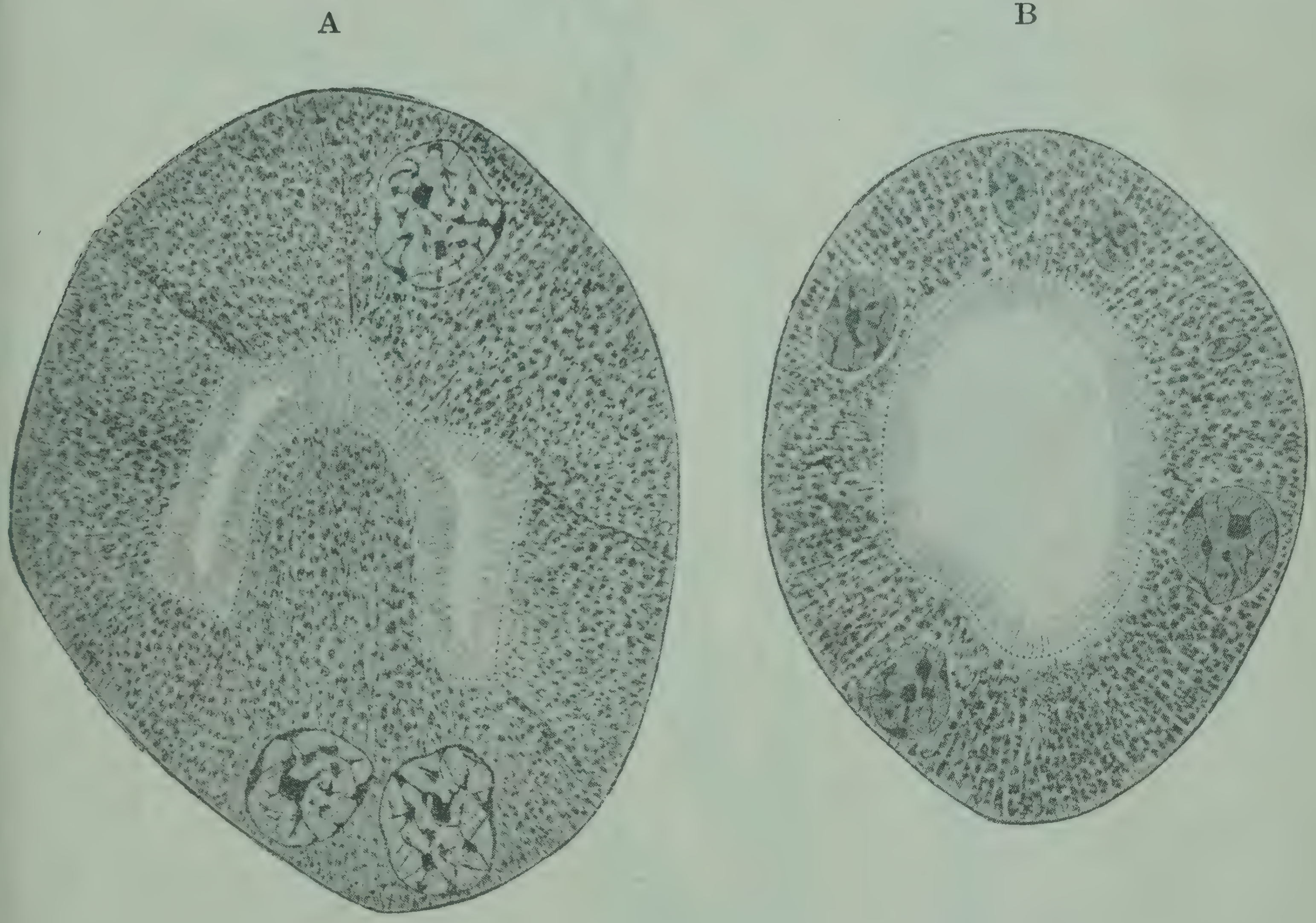


FIG. 607. Cross sections of Convoluted Tubules from Kidney of Rat. (SAUER.)
A. During slight secretion. B. During maximal secretion.

The clear differences between the structure of these various parts point to a differentiation of function. This conclusion is borne out by a study of the blood supply which in the medulla is quite different from that in the cortex. The renal artery divides in the pelvis into five or six branches which pass up to the boundary zone and there give off arteries which run towards the surface and are called the interlobular arteries. Each of these, which is an end-artery presenting no anastomoses with its fellows, gives off on all sides short wide branches, which pass to the glomeruli and constitute the *vasa afferentia* of these bodies. Each vas afferens has a thick muscular wall.

In the vicinity of the afferent arteriole the cells of the distal convoluted tubule are densely packed, and near to the glomerulus the afferent arteriole contains numerous epithelioid cells which are believed to alter in volume, so regulating the blood flow through the glomerulus, perhaps in relation to the contents of the distal tubule; in some conditions of arterial hypertension they become hypertrophied, and are thought by some to be sites of production of renin (p. 659); others believe that renin is produced in the proximal tubules. The glomerulus itself consists of a number of wide capillaries invested by an extremely thin wall. The glomerular capillaries are collected together to form an efferent vessel, the *vas efferens*, which is narrower than the vas afferens and, like the latter, presents a well-marked muscular coat. By alteration in the tone of the afferent and efferent arterioles the blood flow through the cortex and the intraglomerular pressure can be widely varied. The vas efferens breaks up again into a second set of capillaries, which ramify round the tubules of the cortex. From the capillaries of the tubules, the blood is collected again into veins, which leave the kidney partly by the cortex and capsular vessels, chiefly by inter-lobular veins and thence by large trunks which join to form the renal vein at the hilum of the kidney.

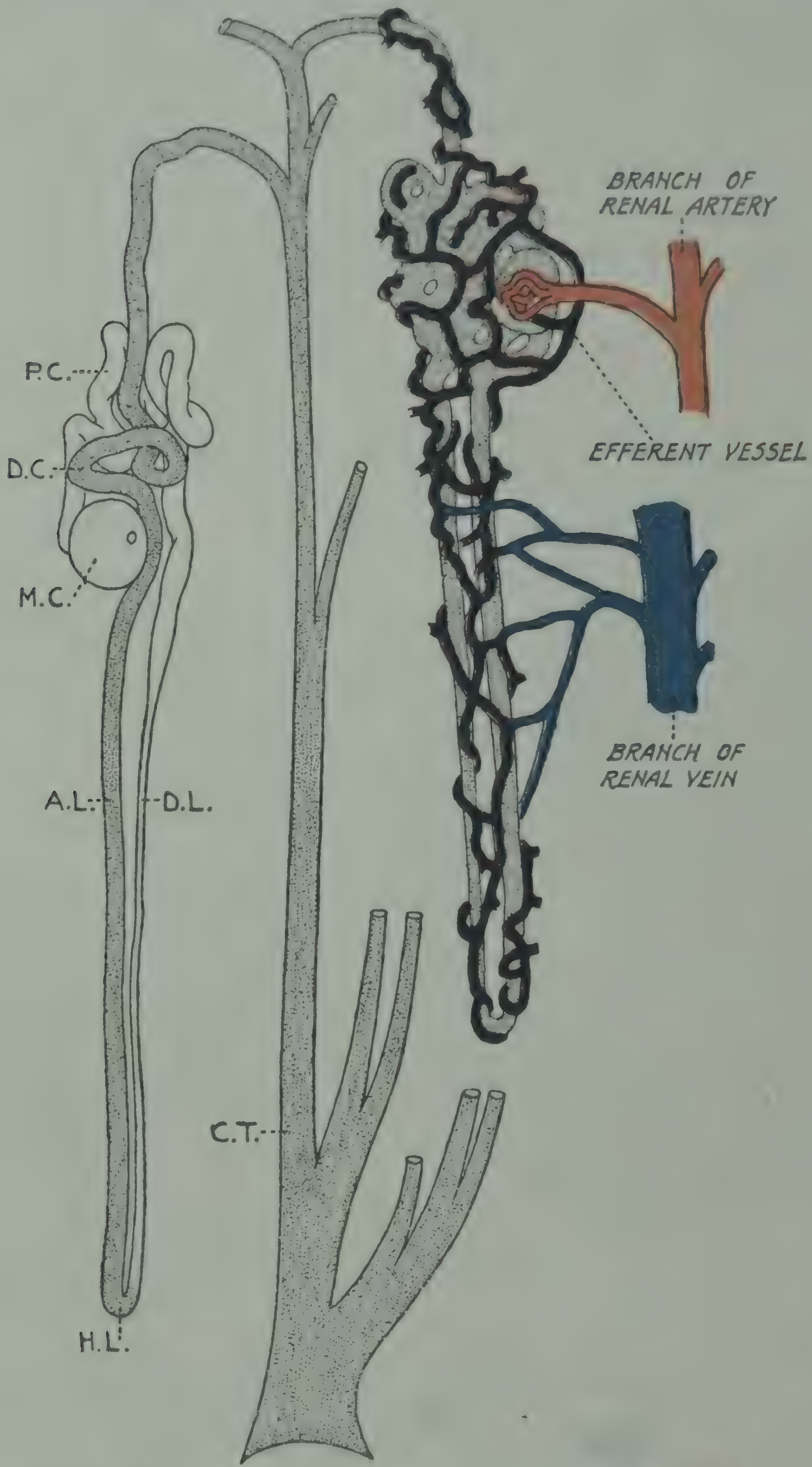


Fig. 608. Diagram of Two Nephrons. On the left, the tubule is outlined from the capsule to the loop of Henle and is shaded from that point to the end of the collecting tubule. On the right, a diagram of the circulation is added. M.C., Malpighian capsule. P.C., Proximal convoluted tubule. D.L., Descending limb. H.L., Loop of Henle. A.L., Ascending limb. D.C., Distal convoluted tubule. C.T., Collecting tubule. (After Cushny.)

The blood supply to the medulla is peculiar, and is derived from the efferent arterioles of the glomeruli nearest to the medulla (juxtamedullary glomeruli); these glomeruli are often simplified in pattern. Their efferent arterioles break up into tufts of long, roughly straight and parallel vessels, which are matched by corresponding veins that open into the inter-lobular veins (Fig. 609). These vessels are called the *vasa recta*, and from the arterioles some capillaries branch off and supply the medullary tubules. But in the main the juxtamedullary arterial vessels loop back directly into the venous ones, at various levels.* This arrangement constitutes an arterio-venous shunt which, when closed, diverts the blood supply to the cortex, and when fully open can divert much of the supply to the medulla and render the cortex ischæmic. Taken together with the vascular control in the cortex itself, this gives a very wide range of separate variability in the blood supply of cortex and medulla.

The kidney is richly supplied with nerves derived mainly from the 6th to 13th dorsal spinal roots,† which are chiefly distributed to the muscular walls of its blood vessels.

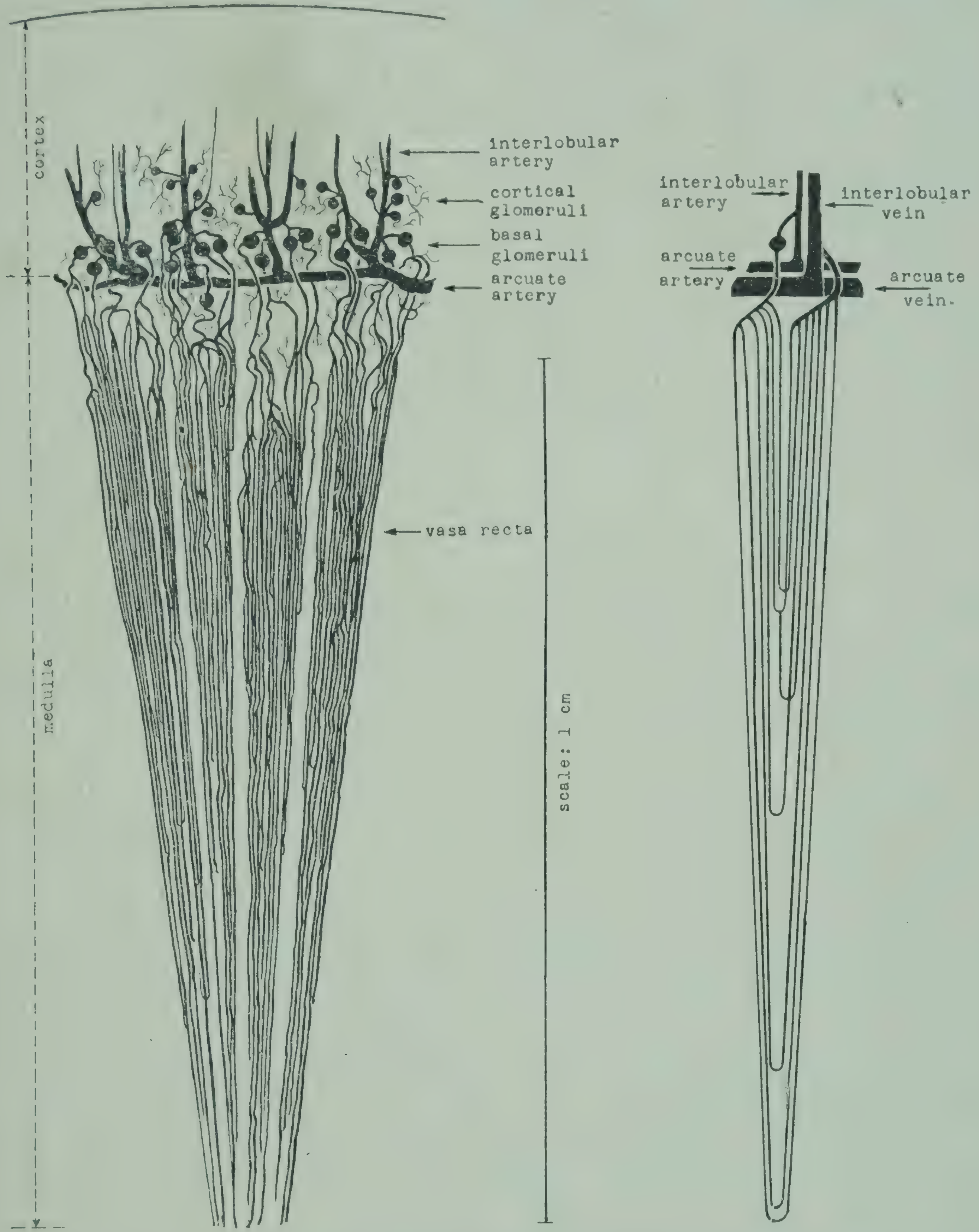

cortex
interlobular artery
cortical glomeruli
basal glomeruli
arcuate artery
vasa recta
medulla
scale: 1 cm
interlobular artery
interlobular vein
arcuate artery
arcuate vein.

FIG. 609. Diagram of blood supply of Cortex and Medulla of the kidney. (Reprinted from "X-ray Studies of the Kidney Circulation" (FRANKLIN, *Science News*, No. 10, 1949), by courtesy of Penguin Books Ltd.)

Some authors have described a fine nerve-plexus surrounding the tubules and sending branches between and into the cells of the convoluted tubules themselves. There is no evidence that the secretion of urine is influenced by the nerves, except through their effects on the blood vessels ‡ : a transplanted kidney functions in agreement with this.§

---

* TRUETA *et al.* *Lancet*, 1946, *ii*, 237. Studies of the Renal Circulation, 1947 (Oxford, Blackwell). FRANKLIN. X-ray Studies of the Kidney Circulation. *Science News*, 1949, No. 10, 101.

† BRADFORD. *J. Physiol.*, 1889, **10**, 358.

‡ PAGE and HEUER. *J. Clin. Investig.*, 1935, **14**, 443.

§ MALUF. *Amer. J. Physiol.*, 1943, **139**, 103.

Relation of Structure to Function. The main points in the structure of the kidney were made out by Bowman * and suggested the two theories of urinary secretion, those of Bowman and of Ludwig (1844),† which have furnished the basis of all subsequent investigation of the subject. Both observers appreciated the great difference between the "membrane" covering the glomerular loop and that of the lining of the tubule, and both drew attention to the difference in the circulation in these two portions of the kidney. The glomerular capillaries, supplied with blood through a short wide artery, and drained by an efferent vessel smaller than the afferent, would represent a region of relatively high capillary blood pressure, whereas the pressure in the capillaries surrounding the tubules must be at least as low as that in most other capillary regions. Bowman therefore suggested that the urine arises from two sources, namely, one part containing the water and salts, produced by a process of filtration through the walls of the glomerular capillaries, and another part containing the specific urinary constituents, urea, uric acid, etc., secreted by the cells of the convoluted tubules and added to the dilute glomerular product as it passed on its course along them.

To Ludwig, on the other hand, it seemed possible, at first, to account for the whole process of formation of urine on the assumption that absorption, and not secretion, was the main function of the tubules. He imagined that the whole of the urinary constituents passed, together with much water, through the walls of the vessels of the glomerulus by filtration from the blood. The glomerular transudate, filtering through what is nowadays called the "glomerular membrane," into the Bowman's capsule, would represent, therefore, an ultrafiltrate of blood plasma, and this view still holds the field as the most likely explanation of the functions of the Malpighian body. Ludwig ascribed the great difference in urea content between the blood and the urine to a concentration taking place in the fluid in its passage along the tubules, in which water and certain of the salts were reabsorbed; a process of reabsorption conditioned by the difference in protein content between the urine within the tubules and the tissue fluid and blood under low pressure on the outside of the tubules.

In its original form, the theory of Ludwig is untenable, because the difference of osmotic pressure between blood plasma and its ultrafiltrate is too small to attract from the filtrate back again into the blood enough water to produce the requisite concentration, even of urea, found in urine.

For the production of urine from the glomerular filtrate, a certain amount of work would have to be done by the cells of the tubules in abstracting the necessary amount of water and returning it to the blood. The minimum work necessary to form a given amount of urine of a given concentration, from blood of average composition, can be found very roughly by comparison of their osmotic pressures, by any of the methods described in Chapter V. For example, we may find their freezing points. A depression of freezing point of 1° C. in an aqueous solution corresponds to an osmotic pressure of 122·7 m. of water and hence, to concentrate a fluid with a freezing-point of −1° C. to one with a freezing-point of −2° C., would require an expenditure of work equivalent to that required to compress 100 c.c. of a gas at a pressure of 122·7 m. of water to half its bulk. In this way it has been calculated that the work necessary to change a fluid of $\Delta = -0{\cdot}56^\circ$ C. (such as plasma) to one of $\Delta = -2{\cdot}3^\circ$ C. (urine), and to yield 1,000 c.c. of the latter would be of the order of 215 kg.m.

More exact calculation would involve the application of thermodynamics, since it would necessitate the separate calculations of the "free energy" changes involved in all the important changes of concentration that occur in converting the glomerular filtrate into the urine. Such calculations have been made,‡ and indicate that for average human urine the kidney does work amounting to about 0·7 gram calories per c.c. of urine formed, or 70 gram cals. per gram of nitrogen excreted. This would amount to about 300 kg.m. per litre of urine, or, taking accepted values for the oxygen usages of the kidneys, say 2–4 c.c./g./hr.,§ an "efficiency" of about 1 per cent. Direct estimation of the constituent of the urine and of the oxygen usages of the kidney also give efficiencies of about 1 per cent for isolated perfused dog's kidneys.‖

---

* Bowman. *Phil. Trans.*, 1842, 57.

† Ludwig. Lehrbuch d. Physiol. des Menschen, 1861, **2**, 373.

‡ Borsook and Winegarden. *Proc. Nat. Acad. Sci.*, 1931, **17**, 13.

§ Fee and Hemingway. *J. Physiol.*, 1928, **65**, 100.

‖ Eggleton, Pappenheimer and Winton. *J. Physiol.*, 1940, **96**, 363.

This amount of work done in order to bring about concentration is not prohibitively high, and in any case is the same, by whatever way the urine is produced. In man, the actual work done is really trivial when compared with that effected by the heart, which, though not weighing much more than the two kidneys, would do at least thirty times as much work in the same time.

However, it is certain that the formation of urine from glomerular filtrate does not occur in any simple fashion. If we compare the composition of blood plasma with that of urine, as in the table given below, we see that the constituents are concentrated to very different extents, so that urine formation, whether by reabsorption or by secretion is a selective process.

RELATIVE COMPOSITIONS OF BLOOD PLASMA AND NORMAL URINE IN MAN.

	Blood plasma, per cent.	Urine, per cent.	Change in concentration in kidney.
Water	90–93	95	—
Proteins, fats and other colloids	7–9	—	—
Glucose	0·1	—	—
Urea	0·03	2	60
Uric acid	0·003	0·05	17
Na	0·32	0·35	1
K	0·02	0·15	7
Ca	0·008	0·015	2
Mg	0·0025	0·006	2
Cl	0·37	0·6	2
PO_4	0·009	0·27	30
SO_4	0·003	0·18	60
Creatinine	0·001	0·10	100

The abandonment of Ludwig's view as to the exact mechanism of the concentration does not, however, place his theory out of court. The questions that will still have to be discussed are whether the chief object of the tubules is the concentration of the fluid produced in the glomeruli, or whether they add to this fluid by a further secretory process, or whether they may not possess both functions, and, in their various parts, alter the fluid flowing through them, either by addition or by withdrawal of water or dissolved constituents. The questions which we have to decide by experiment are :

(1) The nature of the glomerular function, the conditions which determine the amount of fluid formed by the glomeruli, and, especially, whether or not the energy required for the formation of the glomerular fluid is all furnished by the blood pressure within the capillaries.

(2) The function of the tubules, whether they secrete or absorb, and in either case, what substances ; also what parts are played in these processes by the various segments of the tubules, which differ so widely in their histological characters.

FUNCTIONS OF THE GLOMERULI

It is generally agreed that a watery exudation, free from protein, is formed in the glomeruli, and that this becomes concentrated on its way through the tubules, either by the absorption of water and certain salts from it, or by the secretion and addition of urea, uric acid, etc., to it. As to the nature of the glomerular functions, two opinions have been held. According to the Ludwig school, the process is one of simple ultrafiltration, in which, under the pressure of the blood in the glomerular capillaries, some 20 per cent. of the water and crystalloid constituents of the plasma are filtered through the glomerular membrane, leaving the protein constituents behind. This view is now thoroughly vindicated. Heidenhain believed that the process also involved a secretory activity of the glomerular epithelium. If we are to accept the opinion that the only function of the glomeruli is that of filtration, the energy for which is supplied by the blood pressure, certain conditions must be fulfilled. Briefly, they are these :

(1) The fluid passed out from the blood into the Bowman's capsules must be identical with the blood plasma *minus* its colloids.

(2) The blood pressure in the glomerular capillaries must be great enough to form the requisite amount of filtrate, and the amount of filtrate formed must vary with that pressure, and, within certain limits, only with that.

There are other requirements also, which follow in the nature of corollaries to these.

One of the difficulties of experimentation on the kidney is that, according to all reasonable theories, the glomerular filtrate must suffer great modification, in volume, in composition, or in both, as it passes along the tubule. The essential problem then is to attempt to disentangle the functions of the glomerular and the tubular portions, and this has, in part, been solved.

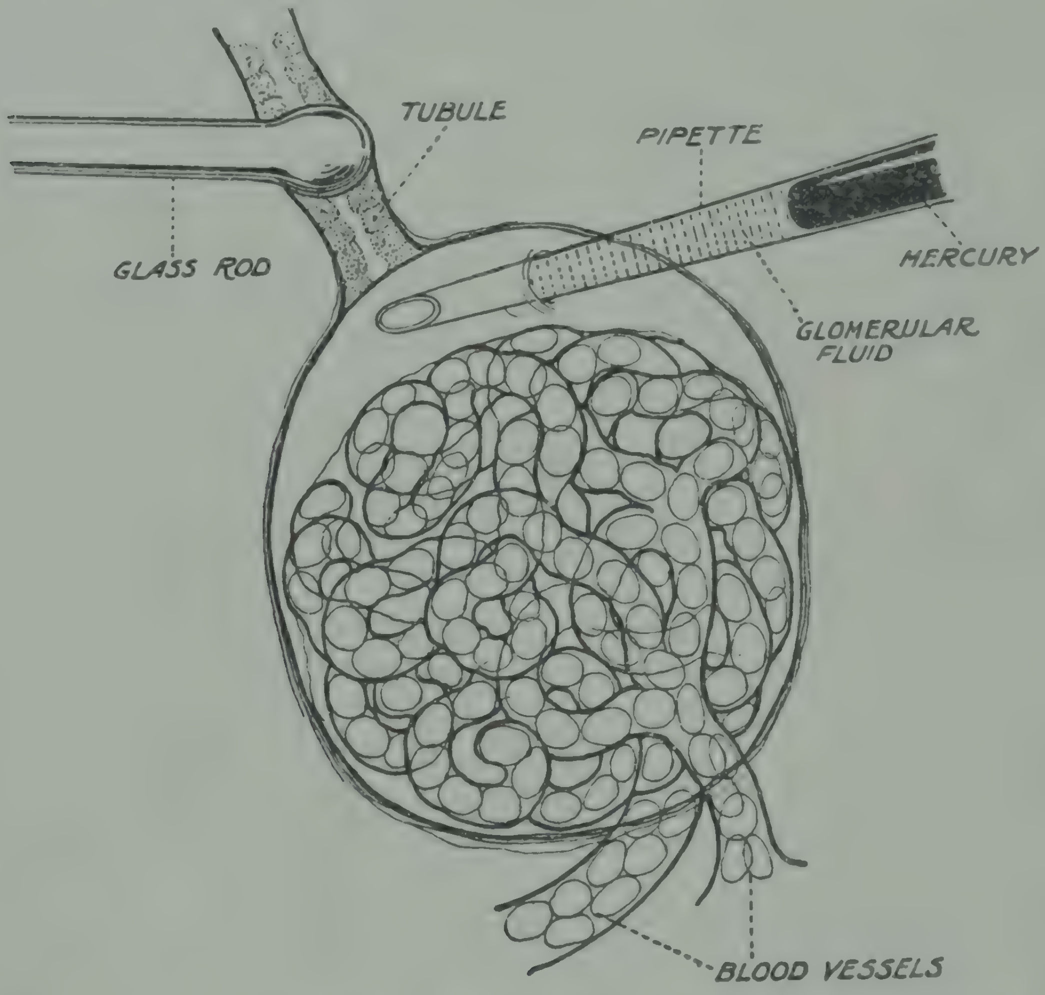


FIG. 610. Diagram to illustrate Richards' method of obtaining samples of glomerular fluid from Bowman's Capsule in the Frog.

The tubule leading from the capsule is closed by pressing on it with a fine glass rod, operated by a micromanipulator (*v.* p. 11). A very fine glass pipette, connected with a reservoir of mercury, is thrust, by a micromanipulator, through the wall of the capsule, and fluid drawn into it by gradually lowering the mercury reservoir. (Winton and Bayliss's "Human Physiology.")

(1) **The Composition of the Glomerular Transudate.** There is abundant evidence that the glomerular transudate has a composition identical with that of the plasma *minus* protein. Thus Richards and his colleagues * inserted capillary pipettes into the Bowman's capsule of the frog's kidney, and so obtained directly a small amount of fluid from the capsule (Fig. 610). By micro-chemical methods they showed that the concentration of chlorides, urea, uric acid, phosphate, glucose or various added substances, such as phenol red, etc., in this fluid, was approximately equal to that in a plasma filtrate from the same blood. The kidney of *Necturus*, which has very large capsules, was similarly examined,† and gave glomerular fluid, which resembled that of the

---

* WEARN and RICHARDS. *J. Biol. Chem.*, 1925, **66**, 247 ; RICHARDS and WALKER. *Ibid.*, 1930, **87**, 479 ; WALKER. *Ibid.*, 499.

† WHITE and SCHMITT. *Science*, 1925, **62**, 334 ; WHITE. *Amer. J. Physiol.*, 1928, **85**, 191.

frog. The fluid in both these animals contained glucose, which is absent from the urine. Further, the general identity of the glomerular fluid with a plasma ultra-filtrate was shown, for both frog and *Necturus*, by Bayliss and Walker * ; they showed that the electrical conductivity of the fluid prepared by filtering the blood plasma through cellophane was identical with that of the glomerular fluid drawn by puncture.

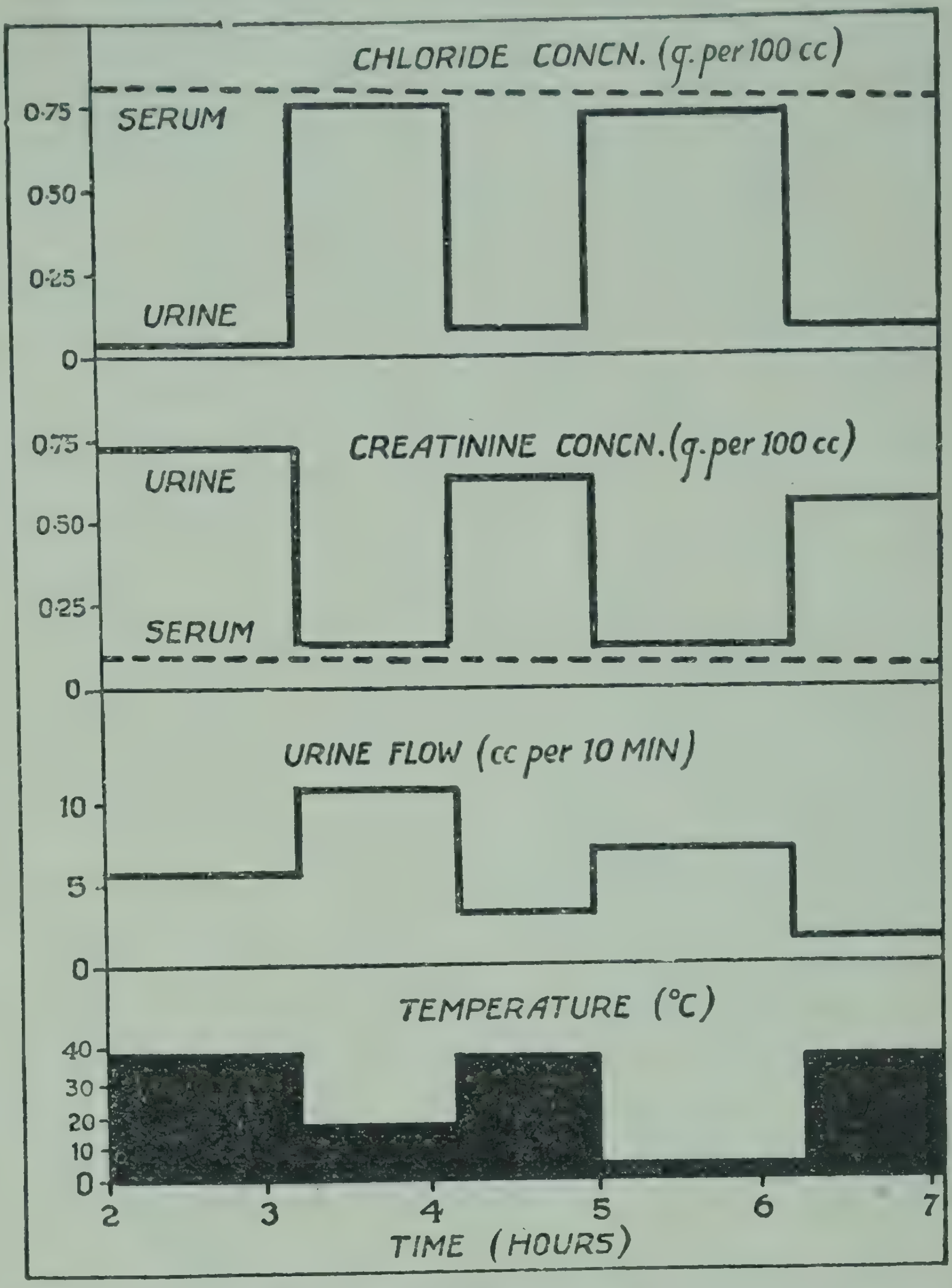


FIG. 611. Effect of Temperature on isolated perfused Dog's Kidney. Cooling put the tubules out of action, so that the urine flow increased, while the chloride content increased and the creatinine content diminished to nearly the same values as in the serum. (BICKFORD and WINTON.)

The collection and analysis of fluid from the glomeruli and other parts of the nephron of mammalian kidneys has also been effected, and with similar results.†

The direct evidence obtained from these experiments is so compelling that there is need for little other.

It has been possible to obtain some indirect evidence on the mammalian kidney. Bainbridge and Evans ‡ connected an excised dog's kidney "in parallel" with the arterial side of a heart-lung circulation, and obtained a secretion of urine. The blood-perfused isolated kidney has since been extensively used by Verney and Starling, and by Winton. A constant circulation of warm oxygenated blood can be maintained, the pressure, velocity and composition of the blood supplied to the kidney being under control. The urine obtained from the perfused kidney contains urea, creatinine, and certain colouring matters in higher percentage than they are present in the circulating blood. The urine is, however, hypotonic and contains only 0·02 to 0·2 per cent. NaCl, as against 0·65 per cent. or more in the plasma of the circulating blood ; creatinine, on the other hand, is present in a concentration several times greater than that in the plasma. It is believed that this is because chloride has been absorbed from the glomerular filtrate as it passes down the tubules, while creatinine has been concentrated by removal of water.

If the fluid produced by the glomeruli is similar in its molecular concentration to the circulating plasma, the alterations on its way down the tubules, however effected, must necessitate the expenditure of energy on the part of the cells lining the tubules. Bickford and Winton § subjected such perfused kidneys to different temperatures and found that when cooled to 25° C. the urine flow was increased, while its composition came very closely to resemble that of an ultra-filtrate of plasma, *i.e.* the chloride increased and the creatinine diminished (Fig. 611). They explained this change as due to the effect of the cold in depressing the activity of the tubule cells and so causing a reduction in the absorption of chloride and of water. A similar effect is produced by poisoning the tubules by adding HCN to the blood: the urinary chloride increases and the urea diminishes.*

---

* BAYLISS and WALKER. *J. Biol. Chem.*, 1930, **87**, 523.

† WALKER *et al.* *Amer. J. Physiol.*, 1941, **134**, 562, 580.

‡ BAINBRIDGE and EVANS. *J. Physiol.*, 1914, **48**, 278.

§ BICKFORD and WINTON. *J. Physiol.* 1933, **78**, 14P.

Further, the composition of the urine approximates more closely to that of glomerular transudate when the filtration is rapid, because the transudate travels more rapidly along the tubules, so allowing less time for alteration to occur in it, the more rapidly the formation takes place. The following experiment shows that this is the case. The rate of secretion of a kidney in a heart-lung-kidney preparation was altered by altering the arterial blood pressure. Sodium chloride and urea were added to the blood.† It will be observed that as the urine flow accelerated the urine chloride rose and the urea fell, both of which changes approximated it in composition to the blood serum.

Time.	Arterial B.P. mm. Hg.	Rate of urine flow, c.c. per 15 mins.	NaCl content, p.c.		Urea content, p.c.	
			Serum.	Urine.	Serum.	Urine.
1·55	111	14·6	0·86	0·35	0·061	0·352
2·22	147	91·5	0·99	0·69	0·048	0·088
2·30	95	43·5	0·99	0·54	0·042	0·112

In accordance with the theory, the total quantity of a solid excreted in a given time should rise with any increase in the urinary flow—for whatever the activity of the tubules, the glomeruli must blindly turn out a constant concentration of solutes. Thus, in the above quoted experiment, the following amounts of urea and chloride were excreted per fifteen minutes:

Urine flow, c.c.	mg. Urea.	mg. NaCl.
14·6	61	51
91·5	81	630
43·5	49	235

Proteins are normally absent from the glomerular filtrate and from the urine, since the glomerular membrane will not allow the normal plasma proteins to pass. It is, however, somewhat permeable, the pores being only just small enough to hold back the plasma proteins, the smallest molecules of which are serum albumin with a molecular weight of 72,000 ‡ and diameter of 38 Å. Hæmoglobin, if present owing to hæmolysis, is passed out and appears in the urine if it exceeds a certain small amount in the plasma; its molecular weight is 67,000. If egg albumin, or Bence-Jones protein, or gelatin (35,000, diam. 18 Å) are added to the blood, they are freely excreted, whereas foreign proteins of high molecular weight, *e.g.* casein (200,000), are not.

Damage to the glomerular membrane by oxygen lack, or by certain diseases, increases the permeability so that the normal serum proteins are excreted, giving rise to the condition of albuminuria. Some persons are at the threshold and pass albumin in the urine when up and about but not when lying in bed (postural proteinuria).

A more likely explanation of postural proteinuria is that the inferior vena cava is compressed against the spine by the liver when the subject is in a lordotic posture, and this causes passive congestion of the kidney and so proteinuria.§

---

* Starling and Verney. *Proc. Roy. Soc.*, 1924, **97B**, 321.

† Starling and Verney. *Proc. Roy. Soc.*, 1924, **97B**, 328.

‡ Bayliss, Kerridge and Russell. *J. Physiol.*, 1933, **77**, 386.

§ Bull. *Clin. Sci.*, 1948, **7**, 77.

(2) **Factors affecting the Amount of Glomerular Filtrate.** The theory that the fluid formed in Bowman's capsule is an ultra-filtrate of the blood in the glomerular capillaries would lead us to anticipate that the rate at which it is formed in each glomerulus would be determined by factors essentially mechanical in nature. These are the ureter pressure, the venous pressure, the glomerular capillary pressure, the intrarenal pressure (*i.e.* the pressure in the kidney substance) and the extrarenal pressure (*i.e.* the pressure on the outside of the kidney). The resultant of all these is called the *effective filtration pressure.* This would be increased by increase of blood pressure in the glomeruli and opposed by increase of pressure in the Bowman's capsule or ureters.

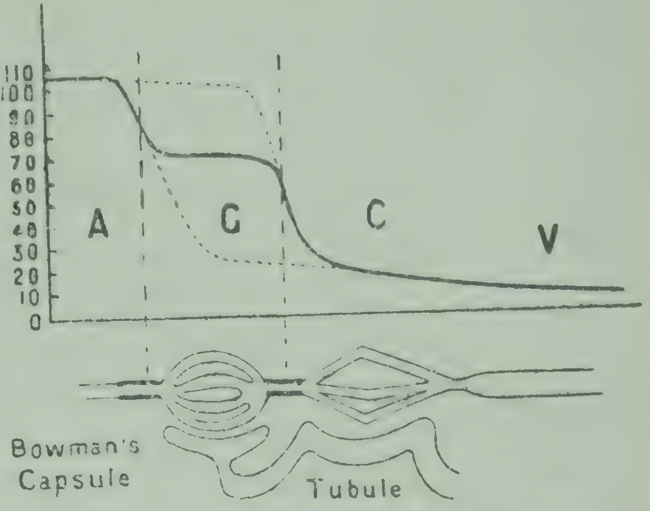


FIG. 612. Diagram (after MORAT) to illustrate the effect of Active Changes in the Vasa Afferentia and Efferentia on the Pressure in the Glomerular Capillaries.

If the vas afferens constricts, the pressure will be represented by the lower dotted line. On the other hand, constriction of the vas efferens would raise the pressure in the glomerulus till it almost equalled that in the renal artery, as is shown by the upper dotted line.

A. Arteries.
C. Tubular capillaries.
G. Glomerular capillaries.
V. Vein.

There are, however, various factors which complicate this apparently simple relationship. One of these is that the number of glomeruli through which blood is circulating is not necessarily constant. It has been shown * that under certain circumstances, glomeruli of the frog's kidney can be seen to constrict so that no blood passes through them. Intermittent closure of the glomeruli is to some extent spontaneous, but the number so closed is greatly increased by stimulation of the kidney nerves, by various reflexes, and by adrenaline. It is thought that closure of the glomeruli in this way is not a frequent phenomenon in the mammalian kidney, and it is now commonly disregarded.

Another complicating factor is possible, should the afferent and efferent glomerular arterioles have different states of tonus, and this condition actually occurs. If a constriction affects the afferent arteriole, the result anticipated would be a reduction of blood flow, a fall in glomerular pressure and a lowered rate of glomerular filtration: on the other hand, with a constriction affecting only the efferent arteriole, the result would be again a reduction of blood flow, but with it a rise of glomerular pressure and of filtration rate (Fig. 612). In the isolated mammalian kidney, adrenaline in very small concentrations causes diuresis with slowing of the blood flow; in larger amounts it lowers the flow of blood still more and then the urine flow is reduced. The smaller doses constrict the efferent arterioles and relax the afferent ones, thus raising the glomerular pressure; larger doses constrict both vessels.† In man, too, adrenaline in small doses acts in a similar manner, reducing blood flow but not filtration rate. In general there is a reciprocal relation between glomerular pressure and blood flow, because in man the renal blood flow is regulated mainly by the efferent arterioles: hence filtration rate tends to remain constant in spite of change in blood flow, if the latter is the result of local regulation.

The pressure in the glomerular capillaries can be raised, not only by an increase of general arterial pressure, but also by a relaxation of the afferent arterioles (if these are constricted), which would diminish the resistance to the blood flow into the kidney. Hence, if, in an anæsthetised animal, in which considerable sympathetic muscular tone exists, the nerves entering the kidney be divided, the arterioles of the organ become dilated and the urine flow is increased.* It should be noted, however, that in the normal, unanæsthetised animal, the renal vessels are apparently not in a state of tonic constriction, since denervation of the kidney is without effect on the rate of urine formation.† In many of the older experiments it was shown that, in general, an increase of the volume of the kidney, due to a rise of arterial pressure, and/or to a relaxation of the renal arterioles, accompanied an increase of urine secretion. Conversely, constriction of the renal arterioles, as by stimulation of the splanchnic or renal nerves, or a fall of blood pressure, led to shrinkage of the kidney and reduction in the urine flow.

---

* RICHARDS and SCHMIDT. *Amer. J. Physiol.*, 1924, **71**, 178.

† WINTON. *J. Physiol.*, 1931, **73**, 151.

An important practical point arises in this connection. It has been stated that reflex suppression of urine secretion may occur after surgical operations; suppression may also follow on serious dehydration due to vomiting, diarrhœa, etc. The probable explanation is that these conditions all decrease glomerular filtration and low blood pressure may be a cause of this; death from renal failure may thus occur without any renal disease.

The urinary secretion can be stopped by obstruction of the ureter. On applying a ligature, the secretion continues for a time, until the pressure in the ureter rises up to a certain point, when the secretion comes to an end. This usually happens when the difference of pressure between urine in the ureter and blood in arteries is between 40 and 50 mm. Hg.

Experiments on the heart-lung-kidney have shown that the effects of alteration of ureter pressure are somewhat complex, but that when the rate of urine formation is changed by alteration of the ureter pressure, the resulting alterations in the composition of the urine are the same as those produced when the rate of urine formation is changed to the same extent by an alteration in the arterial pressure.‡

It was objected by Heidenhain § that, according to the filtration hypothesis, a rise of pressure in the renal vein should raise the glomerular pressure and so accelerate the rate of urine formation, whereas, in fact, it retards it. It appears that this is because the effect of the rise of renal vein pressure produces its greatest effect in pressing upon the tubules and so by increasing the intrarenal pressure, retarding the flow of urine in the same way as does a rise of ureter pressure.‖ Whether the urine flow is retarded by raising ureter pressure, raising the venous pressure, or lowering the arterial pressure, however, the resulting alterations in the composition of the urine are practically identical.

Magnitude of the Filtration Pressure. Working against the mechanical filtration pressure is the colloid osmotic pressure of the blood plasma. The proteins of plasma exert an osmotic pressure of about 30 mm. Hg. Hence, for ultrafiltration to be possible the pressure in Bowman's capsule, and in the ureter, must be at least 30 mm. Hg lower than that of the blood in the glomeruli.

The anatomical arrangements are such as to bring the glomerular blood pressure up to a high value. The vasa afferentia are short, and the vasa efferentia are only two-thirds of the diameter of the vasa afferentia. Moreover, the sudden increase of bed in the bundle of capillaries must itself diminish the fall of pressure in the latter, or might even cause an actual slight rise, owing to the transformation of the kinetic energy of the moving fluid into potential energy, *i.e.*, pressure on the walls of the capillaries. It has been shown by White, however, that this effect is small. In the frog, Hayman ¶ has directly estimated the pressures in the various vessels to be as follows :—

Aorta	37·3 cm. H_2O	
Afferent arteriole of glomerulus .	31·0 ,, ,,	or 84 per cent. of aortic pressure.
Capillaries of glomerulus . .	20·0 ,, ,,	,, 54 ,, ,, ,,
Renal portal vein . . .	5·2 ,, ,,	,, 14 ,, ,, ,,
Renal vein	2·6 ,, ,,	,, 7·1 ,, ,, ,,

---

[*] Cl. Bernard. "Liquides de l'Organism," 1859. Fee. *J. Physiol.*, 1929, **68**, 305.

[†] Klisiecki, Pickford, Rothschild and Verney. *Proc. Roy. Soc.*, 1933, **112B**, 496.

[‡] Winton. *J. Physiol.*, 1931, **71**, 381.

[§] Heidenhain. *Hermann's Hdbch. d. Physiol.*, 1883, **5** (i.), 317.

[‖] Winton. *J. Physiol.*, 1931, **72**, 49.

[¶] Hayman. *Amer. J. Physiol.*, 1927, **79**, 389.

The colloidal osmotic pressure of the frog's plasma was estimated by White to be about 10 cm. H_2O, which would leave a margin of 4 to 10 cm. H_2O for filtration pressure.

An indirect method for estimating the mammalian glomerular capillary pressure * indicates a glomerular pressure about 0·65 of the arterial pressure.

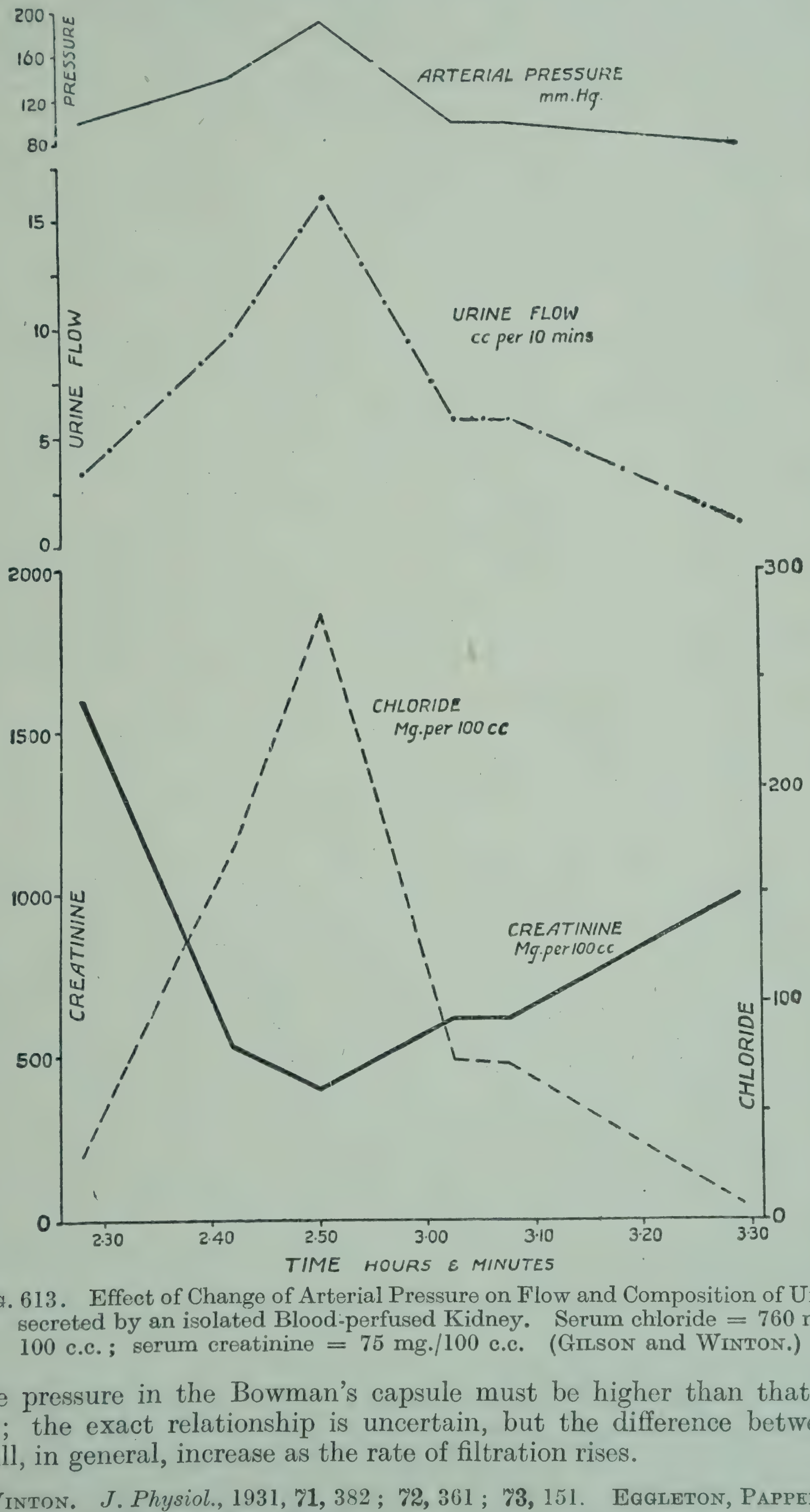


FIG. 613. Effect of Change of Arterial Pressure on Flow and Composition of Urine secreted by an isolated Blood-perfused Kidney. Serum chloride = 760 mg./100 c.c.; serum creatinine = 75 mg./100 c.c. (GILSON and WINTON.)

The pressure in the Bowman's capsule must be higher than that in the ureter; the exact relationship is uncertain, but the difference between the two will, in general, increase as the rate of filtration rises.

---

* WINTON. *J. Physiol.*, 1931, **71,** 382; **72,** 361; **73,** 151. EGGLETON, PAPPENHEIMER and WINTON. *J. Physiol.*, 1940, **99,** 135.

There may thus be, on the two sides of the glomerular membrane, a difference of pressure more than sufficient to cause a filtration of a protein-free fluid from the blood plasma coursing through these capillaries. On lowering the arterial pressure, or on raising the pressure on the tubule side, the filtration ought to come to an end when the pressure in the Bowman's capsule is still 30 to 40 mm. Hg below that in the glomerular capillaries. A number of observers have found that urinary secretion ceases when the arterial blood pressure falls to between 40 and 50 mm. Hg. But if the blood be diluted by the constant infusion of normal saline solution, so as to diminish its protein content, a secretion of urine may be obtained with a blood pressure as low as 18 mm. Hg.

The dependence of the urinary flow on the arterial pressure is shown very clearly in some experiments by Richards and Plant *; they placed a pump in the course of one renal artery of the rabbit, blood clotting having been prevented by hirudin. In this way, the circulation through the kidney was entirely under control, its pressure and velocity being known at any moment. The amount of urine varied directly with the blood pressure and did not depend on the amount of blood flowing through the organ. Similar results have been obtained † in experiments on the heart-lung-kidney preparation (p. 997). The results of such an experiment are shown in Fig. 613, in which it is again shown that when the rate of glomerular filtration is increased by raising the blood pressure, the composition of the urine approaches to that of the serum, *i.e.* chloride increases and creatinine diminishes.

All these considerations strongly support the view that the process occurring in the glomeruli is one of filtration, for which the energy is furnished by the blood pressure. But the glomerular capillaries require oxygen for the maintenance of their physical qualities as a filter. Complete obstruction of the renal artery not only stops the production of urine, but its effects last for some time after the obstruction has been removed, and the first portions of urine obtained when secretion is re-established are found to contain coagulable protein. In the same way, any damage to the glomeruli as a result of defective circulation, of imperfect oxygenation of the blood, or of inflammatory and infective conditions of the kidney, is invariably associated with the passage of protein into the urine.

Clearance Tests. Information of fundamental value, both theoretical and practical, as regards kidney function can be obtained by means of what are called clearance tests. In order to explain these, let us suppose that a given substance, capable of accurate estimation, is naturally present in, or has been artificially added to, the blood. Further, let us suppose that this substance can pass an ultra-filter, such as the glomerular membrane; hence, the glomerular filtrate will contain that substance in the same concentration as it is present in the blood plasma. Lastly, let us suppose that the substance is neither absorbed from nor secreted into the fluid by the uriniferous tubule as the stream passes along it. Then all the substance which was contained in the glomerular filtrate will be quantitatively recoverable from the urine.

Now if the urine, of volume V, formed per minute contained U g. per 100 c.c. of the substance, and if the plasma contained P g. per 100 c.c. of it, then it is clear that $\frac{UV}{P}$ c.c. of plasma have been filtered off through the glomerular membrane per minute and so cleared of the substance. This value $\frac{UV}{P}$ is called the *plasma clearance*, or *glomerular filtration rate* (G.F.R.). It will be evident that all substances which fulfil the above conditions should have the same clearance values; also that substances which are absorbed back into the blood again by the tubules will have lower clearance values, while those which are secreted out from the blood into the tubule will have higher clearance values than the G.F.R.

---

* Richards and Plant. *Amer. J. Physiol.*, 1922, **59**, 144.
† Verney and Starling. *J. Physiol.*, 1922, **56**, 353.

Much work has been devoted to finding substances which fulfilled the above-mentioned requirements and which were either normal plasma constituents, or else were non-toxic. It would appear that the polysaccharide inulin is a suitable substance. Sodium thiosulphate and creatinine have also been used, but in man some creatinine, and in the cat some thiosulphate, is secreted by the tubules and so gives too high a plasma clearance. In the dog, rabbit and sheep, creatinine and inulin give identical results.

When we speak of creatinine clearance, inulin clearance, etc., we mean the volume of plasma which may be regarded as freed from the substance in unit time (*e.g.* c.c. per min.) by the kidneys, as estimated by the employment of the substance named. Since the plasma clearance bears a direct relation to the body surface it is often expressed as $C = \frac{UV}{PA}$, *i.e.* in c.c. per sq. m. of body surface per min., where A = the body surface in sq. m., or as this figure C multiplied by 1·73 (the mean area of an adult in sq. m.).

Investigations by this method, using creatinine, inulin and ferrocyanide, were carried out by Van Slyke and his colleagues [*] on dogs with one kidney removed and the other transplanted to the surface of the body so that blood samples could be taken and the renal blood flow measured.[†] They showed that the actual plasma clearance was about 20 per cent. of the total plasma flowing through the kidney, *i.e.* about one-fifth of the blood plasma was filtered off in the glomeruli. This is called the "*filtration fraction.*" The actual clearance at any time depends, in dogs, and presumably also in other species, mainly on the pressure in the glomeruli. In the normal human male subject under standard conditions the following data have been found by Homer Smith *et al.*[‡] :—

	per 1·73 sq.m. body surface.
Blood flow through kidneys	1,189 c.c./min.
Plasma ,, ,, ,, (= c̄ 58 p.c. of blood)	688 c.c./min.
Inulin clearance (=glomerular filtration rate).	130 c.c./min.
Filtration fraction $\left(= \frac{130}{688} \times 100\right)$ . .	19 per cent.

Women have smaller kidneys than men, and hence the corresponding figures for women are about 84 per cent. of those for men when referred to the same standard area of 1·73 sq. m.

The study of clearances has yielded most valuable information on the functions of the kidneys under normal conditions in the human subject, as the above figures illustrate. In man we can now measure renal blood flow (= Diodrast clearance) (p. 1010), glomerular filtrate (= Inulin clearance), renal tubular mass (Tm_D), and correlate them with plasma concentration, urine flow and concentration, etc. (*v.* also p. 1009).

THE FUNCTIONS OF THE RENAL TUBULES

The glomerular filtrate, formed, as explained above, in Bowman's capsule, must undergo considerable changes on its way through the tubules. These changes might be brought about either by the addition, to a relatively small volume of the filtrate, of a concentrated secretion containing the specific urinary constituents in the required proportions, or else by the reabsorption, from a very large volume of filtrate, of water and various other substances, so that the fully formed urine represents the unabsorbed residue of the filtrate.

---

[*] Van Slyke, Hillier and Miller. *Amer. J. Physiol.*, 1935, **113**, 611.

[†] Rhoads. *Amer. J. Physiol.*, 1934, **109**, 324.

[‡] Smith, Goldring and Chasis. *J. Clin. Invest.*, 1938, **17**, 263.
Homer Smith. *Harvey Lects.*, 1939–40, p. 198.

These two processes, secretion or absorption, would involve the performance of identical amounts of work by the kidney, and there is reason to suppose that both can happen, but that in man absorption is normally the more important.

Absorption in the Tubules. According to Cushny, the whole of the changes by which the glomerular filtrate is transformed into urine may be ascribed to processes of absorption occurring in the tubules. He assumed that the composition of the fluid absorbed is always the same whatever the needs of the organism at the moment. There would thus be but little absorption of urea and none of sulphate, while the whole of the glucose, a portion of the uric acid and the greater part of the sodium, potassium and chloride would be absorbed. According to this view, the constituents of the glomerular filtrate may be divided into two classes, "*threshold substances*" and "*no-threshold substances*," the former being completely reabsorbed when present in the blood in amount below the "threshold" value and hence only excreted in the urine so far as their concentration in the blood exceeds that value, while the no-threshold substances are excreted in proportion to their absolute amount in the plasma, because they are not reabsorbed at all. Thus, of the threshold substances, the glucose of the plasma is normally below the threshold, and it is therefore all reabsorbed and so not present in normal urine. The sodium chloride also comes within the threshold class, but its threshold is usually exceeded in normal conditions, and the excess is then excreted. When the sugar of the plasma rises, as in diabetes, to 0·3 per cent., or when phlorrhizin is given, so as to lower the "renal threshold," glucose appears in the urine and then undergoes concentration, just as urea does. It is doubtful whether there are in man any true no-threshold substances as normal constituents of the blood. The nearest to one appears to be creatinine. Most constituents have thresholds, and others are secreted by the tubules to a small extent.

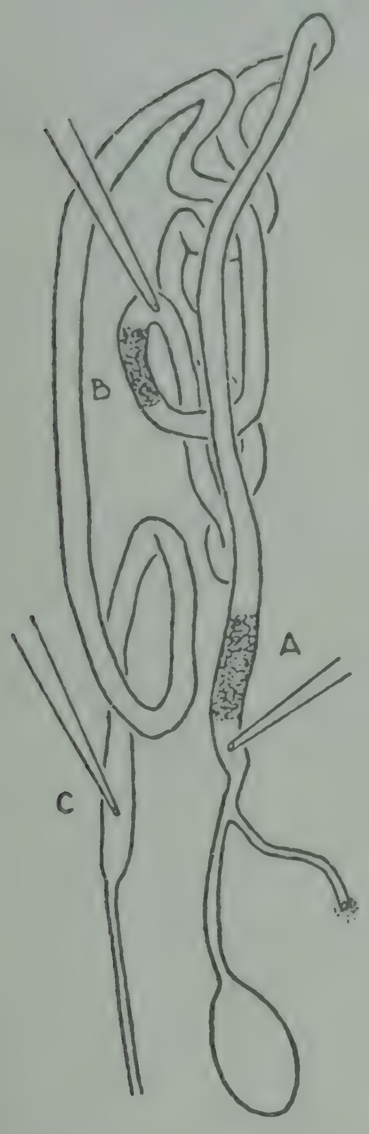


FIG. 614. Insertion of micropipettes for collection of tubule fluid from beginning (A), middle (B), and end (C), of a proximal convoluted tubule of *Necturus*. Stippling at (A) represents blockage of lumen by globule of mercury, at (B) by a globule of coloured oil. (RICHARDS. Croonian Lect., *Proc Roy. Soc.*, 1938, **126** **B**, 398.)

That reabsorption is, in most species of animals, the dominant function of the tubules is now generally conceded. The clearest evidence for it is given by experiments in which it can be directly shown that constituents present in the Bowman's capsule disappear, or are much reduced in amount, during the passage of the fluid along the tubules. We have already learnt that the fluid drawn by puncture from the Bowman's capsule of the frog contains chloride *; yet this substance is absent from the frog's urine. Similar results were obtained with *Necturus* kidneys.† The plasma of the latter animal contains about 0·36 per cent. of NaCl, whereas the urine contains only 0·026 per cent.; when dog's red blood corpuscles were injected into the Bowman's capsule, they were seen to be rapidly washed down into the tubules by the flow of glomerular filtrate; when the corpuscles entered the tubule they were at once laked, presumably because so much chloride had been abstracted from the fluid that it had become hypotonic. Glucose is likewise present in the glomerular filtrate of frog and necturus, but absent from the urine.

Direct proof of the absorptive functions of the tubules was provided by experiments in which the fluid was actually collected and analysed at different points along its course through the tubules. This was effected, in frog and necturus, by the use of capillary pipettes manipulated under the microscope, as shown in Fig. 614.

---

* WEARN and RICHARDS. *J. Biol. Chem.*, 1925, **66**, 247.

† WHITE and SCHMITT. *Amer. J. Physiol.*, 1926, **76**, 483.

Analysis of the fluids collected showed that the chief alterations in total concentration and in Cl occur during the passage along the distal convoluted tubule (Fig. 615), and are probably attributable to reabsorption of both water and solids.

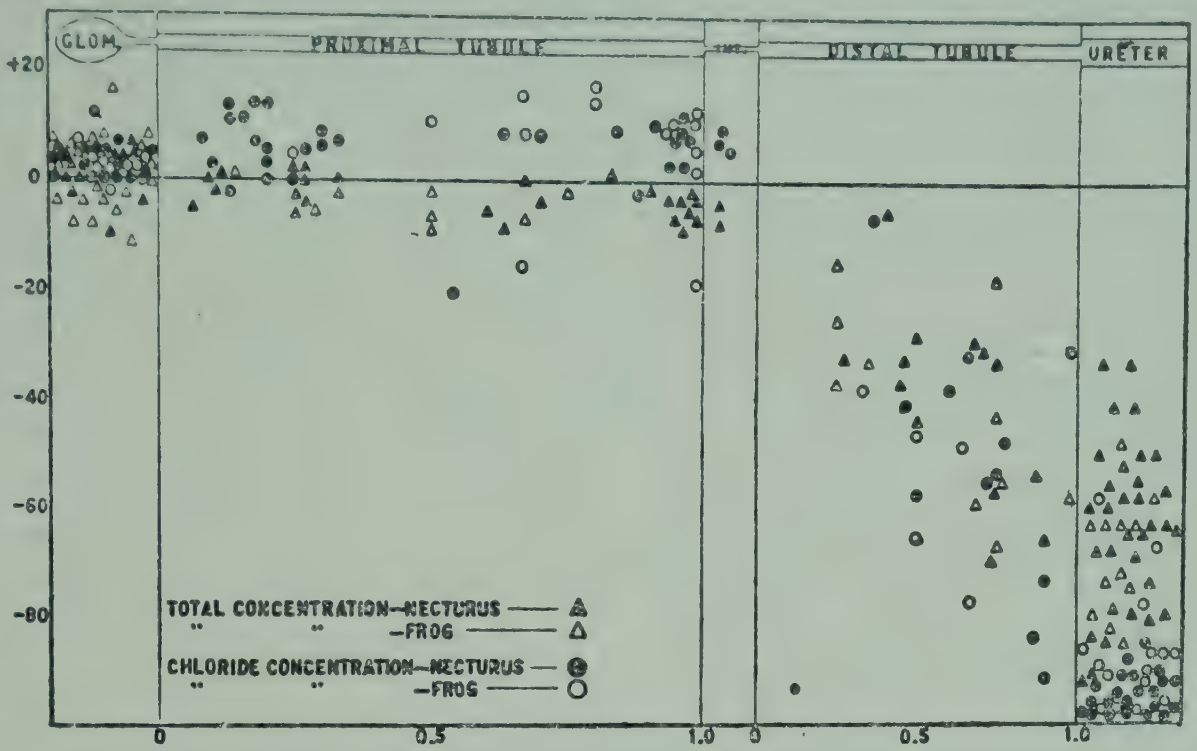


FIG. 615. Chart showing differences between blood plasma and fluid collected from various points on renal tubules, with respect to total solids and Cl concentration. Zero line represents plasma. Cl scale shows per cent. difference from plasma. Concentration of solids represents change in length of column of fluid equilibrated by Barger's method (Osmotic pressure) against plasma. (RICHARDS. Croonian Lect., 1938.)

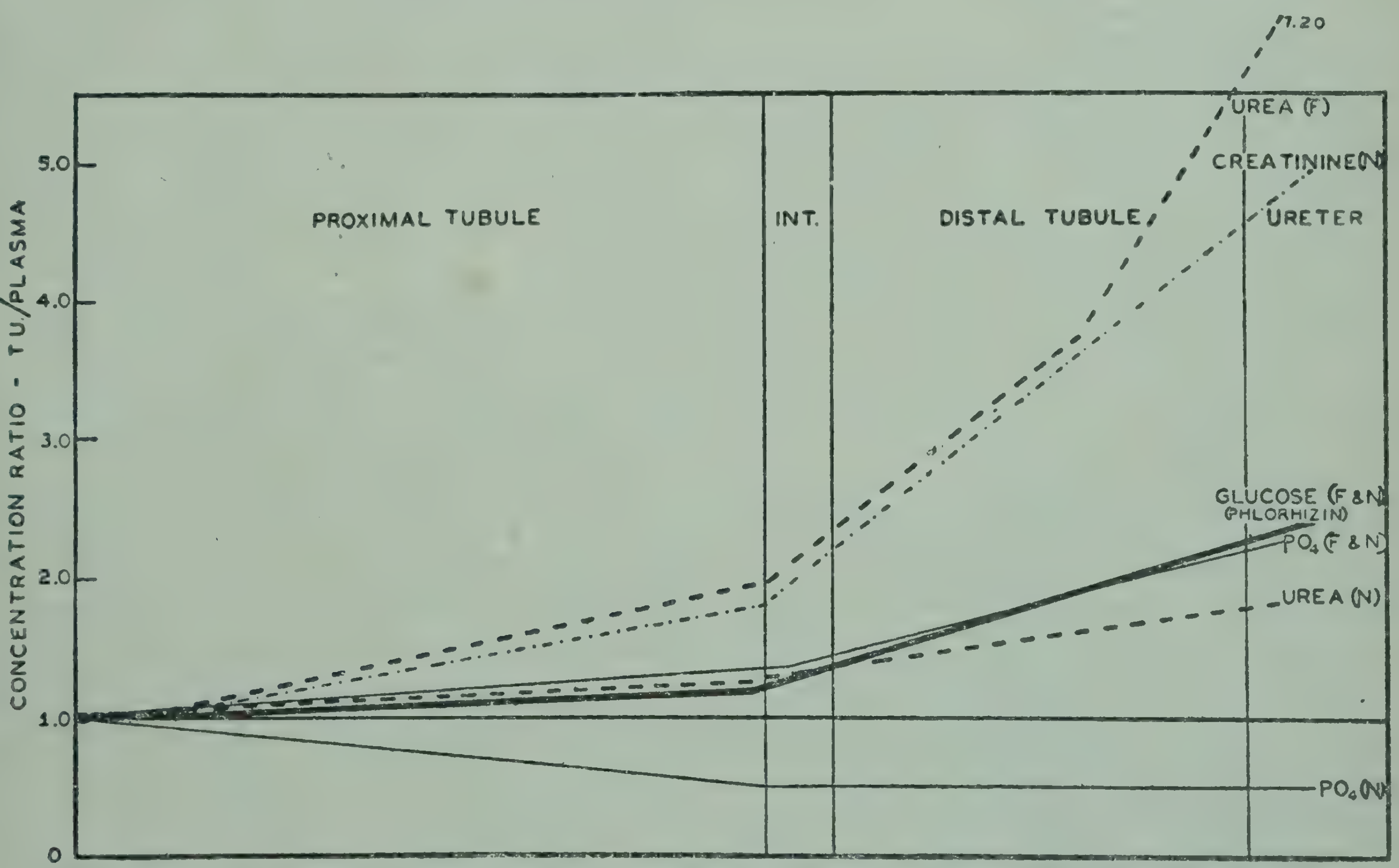


FIG. 616. Average curves of changes in concentration of urea, inorganic phosphate, and creatinine in passing the tubules in frog (F) and *Necturus* (N). Also changes in glucose after phlorrhizin. (RICHARDS. Croonian Lect., *R.S.*, 1938.)

In a similar way, creatinine, urea, glucose and phosphate have been followed. Glucose is absorbed completely, and chiefly in the proximal convoluted tubule ; after phlorrhizin poisoning, the absorption ceases, and there is then some concentration of glucose by loss of water, which occurs in the distal tubule (Fig. 616). As is shown in Fig. 616, inorganic phosphate is also usually concentrated in the distal tubule, but this is again due to absorption of water, and not to secretion. Sometimes phosphate is absorbed back again into the blood ; when this occurs it takes place in the proximal tubule, as is normally the case for Na, Cl, and glucose. Since urea in frog and creatinine in necturus concentrate more rapidly than corresponds with removal of water, it may be that these substances can also be secreted by both the tubules.

There is ample support of other kinds for the belief in the reabsorption of water in the tubules. In the first place, the fluid formed in one Bowman's capsule of the frog can be collected and measured, and when multiplied by the number of capsules, gives the total volume of glomerular filtrate formed in the kidney. Thus Richards calculated that if we take 1 c.mm. per hour for the rate of filtration in one capsule (it is usually higher than this), and the number of capsules in a kidney as about 2,000, most of which are probably functional under the conditions of the experiment, we should get 2 c.c. of glomerular filtrate formed per hour ; but the amount of urine formed per hour under the same conditions is rarely more than 0·2 c.c., so that we must conclude that 90 per cent. of the water passed out into the capsules is reabsorbed in the tubules.

In mammals there is evidence that water absorption occurs not only in the distal convoluted tubule, but also in the thin segment of Henle's loop. Thus pitressin does not exert an anti-diuretic action in amphibia and fishes, which have no thin segment.* It may be, since the thin segment always enters the medulla, that water absorption is passively effected into the vasa recta in the higher animals.†

The older experiments on the excretion of dyestuffs by the kidney were believed to prove that they were eliminated by active tubular secretion. This claim is now very generally admitted.‡ Nevertheless, some of the concentration of dye in the urine can be shown to be explicable by filtration followed by absorption of water. Proof of this has been provided by perfusion of the frog's kidney.§

We may accept, on the balance of evidence, presented here in somewhat biassed form for the sake of simplicity, that there is reabsorption of water and of some dissolved constituents from the glomerular transudate as it passes down the tubules. The dissolved constituents are reabsorbed to different extents ; creatinine little, if at all, and urea not greatly ; chloride, glucose and phosphates more than sulphates.‖

The power of concentration possessed by the cells of the tubules is not indefinitely large, and the urine can therefore never exceed a certain concentration, at which its osmotic pressure just equals the absorptive power of the cells. This limiting concentration differs in different animals, the cat being able to absorb against an osmotic resistance of fifty to sixty atmospheres, while the human kidney cannot concentrate against a resistance of more than twenty-five atmospheres. The presence of any difficultly absorbable substances in the glomerular filtrate, *e.g.* urea, sodium sulphate, or phosphate, must therefore limit the absorption owing to the osmotic resistance they offer to the absorptive powers of the cells. These substances will therefore act as diuretics. In the same way the threshold substances will act as diuretics, when they are present in the plasma in proportions above the threshold, and can no longer be absorbed by the cells of the tubules.

Secretion in the Tubules. Histological examination of the kidney tubules shows a great diversity of structure in the epithelium with which they are lined. Hence, it is quite conceivable that some portions of the tubule might absorb while other portions were capable of secreting : or the same cells might even absorb certain things and secrete others. Yet for one or two decades, physiologists appeared to be so satisfied with the evidence in favour of reabsorption as to forget that the demonstration that urine formation could not be explained wholly on the filtration-secretion hypothesis by no means put that method out of court as an accessory.

---

* Burgess *et al.* *J. Pharmacol.*, 1933, **49**, 237.

† Trueta *et al.* Studies of the Renal Circulation, 1947. (Blackwell.)

‡ Marshall and Vickers. *Johns Hopkins Hosp. Bull.*, 1923, **34**, 1.

§ Richards and Barnwell. *Proc. Roy. Soc.*, 1928, **102B**, 72 ; Richards. "Methods and Results," etc., 1929, p. 34.

‖ Rehberg. *Bioch. J.*, 1926, **20**, 447, 461.

Once an instance had been produced of urine formation in which secretion indubitably played a part, the door was open to admit secretion as a possible collaborator with absorption, in the modification of the glomerular filtrate. Such an instance was produced when Edwards, Marshall and Homer Smith,* recalling the fact that certain fish had kidneys from which glomeruli were absent, pointed out that urine formation must in those species be a purely secretory phenomenon. Their further investigations on these aglomerular kidneys of at least twenty-five species of fish belonging to five different families have, more than anything else, provided the faith that was required for acceptance of the possibility of tubular secretion.

It is suggested that the explanation of the differences in kidney structure in different species of animals is bound up with their phylogenetic history, the glomerulus being developed when water was abundant, and reabsorption then being necessary to prevent loss of valuable solutes ; when water was scarce, or the animal living in hypertonic sea-water, kidney function could be carried out either by secretion alone, the glomeruli being lost, or by reabsorption with or without the addition of secretory power, the glomeruli being retained.

When the possibility of secretion by the tubules had been demonstrated, it was natural to re-examine the classical experiments of the 'seventies, which had once been thought to provide evidence in its favour, and which had in the meantime been explained in other ways. Heidenhain,† for instance, attempted to throw light on the excretory functions of the kidney by studying the mechanism by which it excretes sodium sulphindigotate (" indigo carmine.") If the dye be injected intravenously, it is excreted in a concentrated form by the kidney, which is dyed blue, especially in the medulla.

It has been objected to these results that they may equally well be brought about by absorption. If indigo carmine is turned out by the glomerulus, it will be in so dilute a solution that the glomerulus will not be stained. As the transudate descends the tubules it undergoes concentration, leading to precipitation of the dyestuff in the lumen of the tubules, and the granules in the cells may be the result, not of secretion but of absorption. Some experiments by Marshall, however, certainly point to the active secretion of phenol red by the dog's kidney. Phenol red is partly held in the plasma in combination with protein, and hence is only partly separable by ultra-filtration, 75 per cent. of it remaining behind. Yet the rate at which phenol red passed out in the urine was so large that in order to explain it by reabsorption, the blood flow through the kidneys would have to be over 1 l. per minute, even if half the plasma filtered through and was reabsorbed. Hence the phenol red must have been torn away from its colloidal union, either by secretion in the tubules, or, which is hardly likely, by active secretion at the glomerular membrane.

By taking advantage of the peculiarities of the renal circulation in the frog, claims of another kind have been made for the possession of secretory functions by the tubules. A number of renal arteries supply the glomeruli, the vasa efferentia from which pass, as in the mammalian kidney, into the intertubular capillaries. These, however, also have connections with the renal portal vein, shown in Fig. 617, and it is possible, by perfusing the renal portal vein at a high pressure, to cause a flow in that direction. If all the renal arteries be ligatured, the glomeruli were stated ‡ to have been entirely cut out of the circulation, though the tubules still receive venous blood through the renal portal vein. Ligature of all the renal

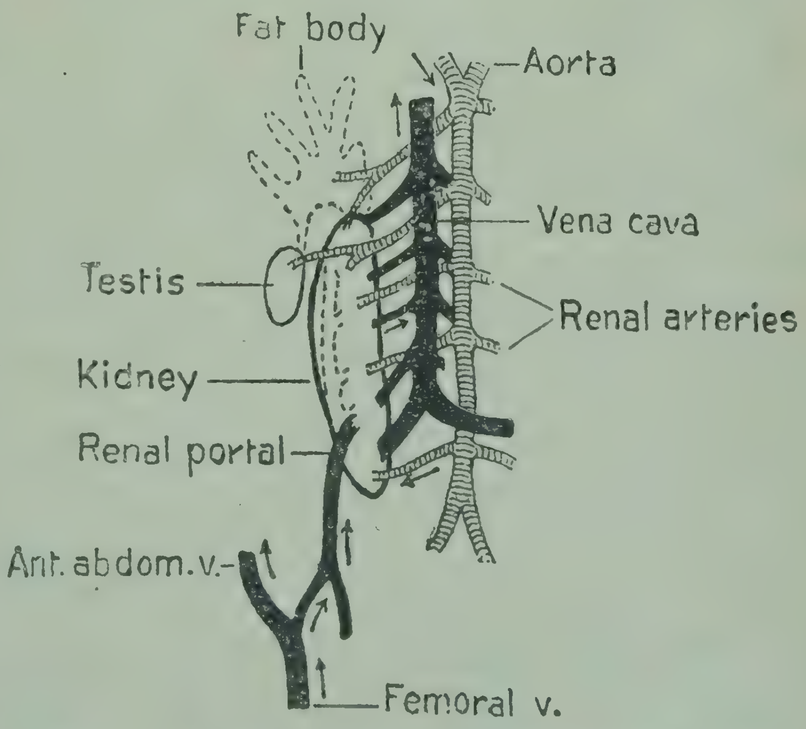


Fig. 617. The Vascular Supply to the Kidney in the Frog.

---

* Edwards. *Anat. Rec.*, 1925, **44,** 15 ; Marshall. *Bull. Johns Hopkins Hosp.*, 1929, **45,** 95 ; Marshall and Homer Smith. *Biol. Bull.*, 1930, **59,** 135 ; Marshall. *Amer. J. Physiol.*, 1930, **94,** 1.

† Heidenhain. *Pflüger's Arch.*, 1875, **9,** 1 ; *Arch. mikr. Anat.*, 1874, **10,** 30.

‡ Nussbaum. *Pflüger's Arch.*, 1878, **16,** 139 ; **17,** 580.

arteries caused cessation of urinary secretion, which could be reinduced by injection of urea, *i.e.*, urea, with water, was secreted by the tubules. This experiment has often been repeated,[*] and with similar results. There has been some difference of opinion as to the accuracy of the anatomical basis of them. Hayman,[†] and Bensley and Steen,[‡] however, consider that Nussbaum's original description is substantially correct.

Experiments by Bensley and Steen on the excretion of dyes by the frog's kidney, suggest that these are actually secreted by the cells of the proximal convoluted tubule, and that there is absorption of water, and perhaps of threshold substances, in the distal convoluted tubule. In these experiments, the kidneys were observed under the microscope at intervals after the injection of indigo carmine or phenol red. Some of the renal arteries were tied, and their nephrons were observed to have no glomeruli functioning. In these, the dye was seen to appear in the cells of the proximal tubule in vacuoles, *which then passed out* into the lumen of the tubule. When indigo carmine was given first, and phenol red after it, the first dye was displaced inwards into the lumen by the second, which followed it along the whole tube, both becoming concentrated as the distal tubule was traversed. Similar results have been obtained by observation, under the microscope, of frog's kidneys with glomeruli put out of action by ligature of the renal artery, and on rat's kidneys in which glomerular circulation was stopped by section of the spinal cord. The dye used was fluorescein, which was rendered visible by ultra-violet light.[§]

To sum up, it seems probable that secretion, when it occurs, is effected by the proximal convoluted tubule, which is also able to absorb some water and certain other constituents, such as glucose and salts. Concentration by withdrawal of water is effected chiefly by the distal convoluted tubule, as is also, probably, the formation of ammonia.

Secretion of the usual urinary constituents does not seem to occur to any great extent in the mammal, though in man some creatinine, and in birds and reptiles uric acid, are definitely excreted. Unusual substances, such as dyes, diodrast (organic iodine compound used for X-ray study of the kidneys), *p*-aminohippuric acid, penicillin, etc., are certainly excreted by mammalian tubules.

Urea Clearance. We have seen that plasma clearance may be used as a means of assessing the volume of glomerular filtrate, provided that the substance chosen for estimation is neither absorbed nor excreted by the tubules. As measured in this way, inulin gives clearances of the order of 123 c.c. per minute. Urea, on the other hand, gives values of only about 75 c.c. per min., so that there is apparently some reabsorption of this substance. If the urea level of the blood is raised by taking urea by the mouth the clearance is unaltered, *i.e.* the same fraction of the urea is removed. But if the glomerular filtration rate is reduced by reduction of the blood flow through the kidney, or by disease, the clearance is diminished. The urea clearance is thus a convenient practical index of the rate of glomerular filtration: it does not actually give us this, as inulin or creatinine clearance does, but it is proportional to it. In nephritis, urea clearance may be only 4 or 5 c.c. per minute.

To carry out a urea clearance test the subject drinks about a pint of water in order to produce a good diuresis. The bladder is emptied to begin with, and again after one hour and two hours, and the volume V in c.c./min. is noted. A blood sample is taken at one hour, and the urea content of blood (B) and urine (U) are determined. The clearance $C_{\text{ur.}} = \frac{UV}{B}$ c.c./min. or, in percentage of the normal clearance,

$$C_{\text{ur.p.c.}} = 1{\cdot}33\,\frac{UV}{B}$$

If the urine flow is less than 2 c.c./min. the formula is:

$$C_{\text{ur.p.c.}} = 1{\cdot}85\,\frac{U\sqrt{V}}{B}.\|$$

---

[*] Bainbridge and Beddard. *Bioch. J.*, 1906, **1**, 255; Atkinson, Clark and Menzies, *J. Physiol.*, 1921, **55**, 253.

[†] Hayman. *Amer. J. Physiol.*, 1928, **86**, 331.

[‡] Bensley and Steen. *Amer. J. Anat.*, 1928, **41**, 75.

[§] Ellinger. *J. Physiol.*, 1940, **97**, 433.

[‖] Peters and Slyke. Quantitative Clinical Chemistry, II., 1932, p. 564. (Williams and Wilkins, Baltimore.)

The normal range is from 75 to 130 per cent. of the average value of 75 c.c./min.

Instead of the urea clearance, we can determine the $\frac{\text{Urine}}{\text{Blood}}$ urea concentration ratio, using morning urine. This is normally about 50.

Various functional tests for the efficiency of the kidney have been proposed for clinical use. Most of them depend on the determination of the degree of concentration in the urine of substances present in or added to the blood, or of clearances of various constituents. Books on clinical chemistry should be consulted for an account of these tests.*

Rate of Tubular Transfer. If a given substance is absorbed, or secreted, by the tubules, the rate at which this transference happens cannot exceed a certain maximum value.† This value, in mg./min., is called T*m* for the given substance and kidney, and is constant. It is estimated, in the case of absorption, by finding the plasma concentration of the given substance, the rate of glomerular filtration (inulin clearance), and the amount of the substance passed out by the urine in a given time. The product of the first two gives the "filtered load"; this is raised in stages by increasing the plasma concentration, by injections of the substance concerned, the excretion rate at each stage being measured by urine analysis. Deduction of the excreted from the filtered load gives the amount of the substance transported, *i.e.* in this case, absorbed back into the blood. With rising plasma concentration or filtration rate, this amount increases from an amount practically equal to the filtered load until it reaches a maximum, beyond which further increase in the plasma concentration produces no further absorption, the additional amount filtered being all passed on into the urine. This maximum value is the T*m* for the substance and individual at the time.

The following example will illustrate :—

GLUCOSE PLASMA CONCENTRATION AND RENAL TUBULAR ABSORPTION (Dog, 24 kg.)
(Shannon and Fisher.)

Period, mins.	Plasma Glucose mg./100 cc.	Glucose Filtered mg./min.	Glucose Excreted mg./min.	Glucose Reabsorbed mg./min.
10	128	150	0·2	150
10	238	267	2·2	265
10	437	499	229·0	270
8	1234	1494	1212·0	282

(In the interval between each period and the next, glucose was injected intravenously.)

It will be seen that, as the plasma glucose is raised, the glucose reabsorbed rises until it reaches the maximum rate (T*m*), in this case about 270 mg./min. : the amount excreted, at first negligible, increases rapidly as soon as the filtered load exceeds the value T*m*, and thereafter all above this is passed out.

Maximal rates of absorption have been demonstrated to occur with several substances, *e.g.* for glucose and *p*-aminohippuric acid (T*m* for P.A.H. = 77·5 ± 13 mg./min. in man), ascorbic acid, sulphate, inorganic phosphate, etc., and the case probably applies to threshold substances in general.

A similar relationship holds good for substances which pass in the other direction, *i.e.* are secreted from the blood into the tubule, a maximal rate being reached when the plasma concentration and flow reach a certain figure. If diodrast, for instance, is given in sufficiently large amount, the T*m* for diodrast ($\text{T}m_{\text{D}}$) can be obtained ; it is nearly independent of the blood flow through the kidneys, and can be used as a measure of the amount of functional proximal tubule tissue (tubular excretory mass), *i.e.* of the size of (normal) kidneys.* It is often convenient to refer the other clearances, *e.g.* inulin (= glomerular filtration) to Tm_p (*i.e.* to kidney size) rather than to body-surface. When we do so we find that whereas in normal subjects inulin clearance ranges from 90–180 c.c./min./1·73 sq. m., the value, related to Tm_p, is almost constant at 2·6 c.c./min./unit Tm_p, in women as in men. The value of Tm_p is a measure of the mass of active tubule tissue, and its constancy shows that all tubules are probably in action all the time.

---

* HARRISON. "Chemical Methods in Clinical Medicine," 1947, p. 86. (Churchill ; London.)

† SHANNON and FISHER. *Amer. J. Physiol.*, 1938, **122**, 765.

These results are explicable on the hypothesis that the substance being transported across the tubular epithelium combines temporarily with a constituent of the cell, and that the limit to the rate at which this takes place is fixed by the laws of mass action and reaction velocity. Phlorrhizin, for instance, probably blocks the combination of glucose and so stops reabsorption.

It will be seen that the view here taken is the opposite one to that of threshold substance propounded by Cushny; he places the barrier to further absorption at the plasma concentration; the view of Shannon is that the determining factor is the absorbing capacity of the renal cells, as fixed ultimately by their chemical content.

Renal Blood Flow. Diodrast, or alternatively *p*-aminohippuric acid, is excreted so vigorously by the tubules as to be completely removed from the blood flowing through the kidney if the concentration in the plasma is low. Hence its plasma clearance will be equal to the flow of plasma through the kidneys, so that if the plasma content of the blood is also known the blood flow can be deduced. It has been found in this way that the plasma clearance for P.A.H. = 621 ± 133 c.c./min./1·73 sq. m., and hence the blood circulation through the kidneys of an average man of surface area of 1·73 sq. m. is about 1,000 c.c./min. or about 20 per cent. of the cardiac output.

THE ADAPTATION OF THE RENAL FUNCTIONS

When we examine the urinary secretion in the intact animal, we cannot but be struck by the manner in which its amount and concentration are adapted to the needs of the body as a whole. The behaviour of the tubule cells with regard to any given substance must therefore depend to a certain extent on the previous nutritive history of the body.

When a large quantity of sodium chloride is taken, the result is a temporary rise in the salt concentration of the plasma which may last a day or two. This is then succeeded by an increased excretion of salt by the kidney, but this occurs at a period after the salt concentration of the plasma has returned to normal. The salt ingested seems at first to pass into the tissues, and is only got rid of from these by an increased excretion after a considerable latent period. Moreover, when diuresis is produced in such an animal by the injection of equivalent quantities of sodium chloride and sodium sulphate, there is no diminution of the NaCl in the urine towards the end of the diuresis, but its percentage rises steadily as the rate of urinary flow diminishes. On the other hand, a total deprivation of sodium chloride extending over several days, although not altering to any large extent the percentage amount of this salt in the blood plasma, leads to a total disappearance of the salt from the urine, the whole of the sodium chloride present in the glomerular filtrate being now reabsorbed in the tubules. When large quantities of water are taken by the mouth, there is a rapid excretion of a very dilute urine, so that the constitution of the body, and of its fluids, is but little altered by the water ingested.

The endocrine organs play a considerable part in regulating the electrolyte and water balance of the body, and this they do in part by an effect on the kidney functions. Conspicuous among these actions is that of some adrenal cortical hormones, and of the steroids related to them, the most active being desoxycorticosterone and progestin. These hormones increase the absorption of Na and Cl in the proximal tubules, and so aid their retention in the body. This action is opposed by that of the antidiuretic hormone of the pituitary, which impairs electrolyte retention and augments water reabsorption; and is favoured by the thyroid hormone, which reduces water reabsorption.

The kidney plays an important part in the regulation of the hydrogen ion concentration of the blood. When there is increased production of acids in the body the excess of acid is excreted in the urine. Normally, fixed acid production exceeds available base increase by 50–100 M.Eq. per day.

---

* Smith *et al.* *J. Clin. Inv.*, 1938, 17, 263.

Three factors appear to control the rate of acid excretion by the kidney, viz. : the amount of buffer in the glomerular filtrate, the nature of that buffer, and the plasma alkali reserve (not its *p*H). The excretion of acid is assisted by the formation of ammonia (p. 932) in the kidney, as well as by alteration in the ratio of mono- to di-basic phosphate in the urine.

The formation of acid urine can only in part be explained by reabsorption of bicarbonate or phosphate. It is probably effected, in the distal tubule, by passage of H-ions into the tubule in exchange for Na-ions which leave it. In this exchange carbonic anhydrase seems to play a part,* as it does in acid secretion in the stomach : if the enzyme is poisoned by sulphanilamide, alkaline urine is formed.

Normal urine contains no sugar, but if the sugar in the blood rises to 0·2 per cent., the excess of sugar is at once excreted. The simplest explanation of the varying activity of the kidney would be to assume that this organ reacts to very minute changes in the composition of the circulating blood. For instance, Clark † perfused the arterial system of the frog with Ringer's solution containing 0·1 per cent. of glucose, and at the same time supplied the tubules *viâ* the renal portal vein with Ringer's solution of varying glucose content. When this solution contained up to 0·1 per cent. of glucose, all the glucose passed out into the Bowman's capsule was reabsorbed, so that none was present in the urine ; when it contained more than 0·1 per cent. there was less and less reabsorption as the amount supplied to the tubules was increased.

The kidney by itself has much less latitude in its activities ; nervous control can apparently be discounted as a party to such collaborative regulation, but the humoral control, through the endocrine organs, is important.

It should be noted that the kidney of the new-born infant is a relatively inefficient organ : the urine it forms is hypotonic, and clearances of inulin, urea and chloride are low : hence the importance of controlling protein, water and salt intake, and of avoiding saline infusions.‡

Adaptation to Reduction of Kidney Tissue. After removal of one kidney, the blood flow through the other one is increased by 70–100 per cent., but it is not known how this result is caused : the urine flow from the remaining kidney also increases, to be about equal to that previously produced by both : the remaining kidney soon shows hypertrophy. The adaptability of the kidney in performing its functions is well seen by the effect of removal of a portion of the kidney, or of deprivation of the blood supply to part of it by ligature of one of the branches of the renal artery. In either case the kidney thus reduced secretes no less urine than the other kidney, which remains intact.§ There is thus a considerable margin of reserve in the kidney. It has been suggested that some of this reserve is provided by the opening-up of glomeruli which had previously been closed down by constriction of their afferent arterioles. Subsequent work, however, discourages this opinion, and regards the intermittent functioning of the glomeruli as being confined to the cold-blooded kidney.

When kidney tissue is considerably reduced, as, for example, by removal of one kidney and half the other, the volume of urine excreted may even be increased,‖ the urine being of a dilute type more nearly representing the glomerular transudate. These results have been obtained also on the heart-lung-kidney preparation,¶ and are probably due to the effect of reduced intrarenal pressure.* The increased urine production of the remaining unit is accompanied by their failure to concentrate the urine formed, because of it more rapid transit along a relatively small number of tubules.

---

* Pitts and Alexander. *Amer. J. Physiol.*, 1944, **142**, 648 ; 1945, **144**, 239.

† Clark. *J. Physiol.*, 1922, **56**, 201.

‡ McCance and Young. *J. Physiol.*, 1941, **99**, 265.

§ Verney. *Lancet*, 1930, (ii), p. 63.

‖ Bradford. *J. Physiol.*, 1899, **23**, 415.

¶ Verney. *Lancet*, 1929, March 30th, July 12th.

A similar condition is seen in pathological states where there has bee destruction of a large amount of kidney tissue, and Verney has appropriatel compared this with the dilatation seen in a failing heart during exercise ; th remaining kidney tubules have, so to speak, become worked to the limit o their functional reserve. We therefore have two types of reserve in the kidne tissue—structural reserve by the occasional closing down of those glomerulo-tubular units whose services can be temporarily dispensed with, and, secondly, the functional reserve represented by the extent to which the tubules, in effecting their work of concentration, are working to the limits of their capacity.

DIURESIS

An increase in the rate of urine formation is called diuresis. It may be caused by copious water drinking, by pathological states, or by the action of alcohol, or of certain drugs.

Two types of fundamental factor might be concerned in promoting an increased flow of urine. These are :

(1) Increased glomerular filtration from mechanical causes, *e.g.* rise of pressure in the glomerular capillaries, reduction of protein content of the blood, or (hypothetically) by the opening-up of glomeruli previously closed by constriction.

(2) Reduced absorption in the tubules, such as may conceivably happen when the diuretic substance is not absorbed, or when the water absorption by the tubules is impaired, as when the post-pituitary hormone is reduced.

In fact, diuresis as commonly met with in man is due to the second of these causes.

Clues to the way in which diuresis is brought about are provided by the study of the osmotic work and oxygen usage, and so, from these, of the efficiency of the isolated kidney.

Pressure Diuresis. When diuresis is caused by an increase of arterial pressure, it would be expected that if the diuresis was only due to increased glomerular filtration, and not at all to altered tubular absorption, there would be no increase in the oxygen usage of the kidney, and no change in its efficiency. Actually, however, a rise of arterial pressure does cause an increase, not only in the filtration (creatinine clearance), but also in oxygen usage, osmotic work and efficiency.† Doubling the urine flow in this manner causes about 40 per cent. increase in osmotic work and 25 per cent. in efficiency. We may infer, therefore, that even in this apparently simple case of increased glomerular activity, there is a concomitant *increase* in tubular reabsorption. The concentrations of all the constituents of the urine during pressure diuresis move towards their concentration in the plasma, so that the work done per unit volume of urine produced is less, but owing to increased work on the part of the tubules, the work done per unit of time is raised. These facts were discovered by the use of the isolated kidney.

Dilution Diuresis. When the plasma colloids are diluted by the addition to the blood of Ringer's solution, the effective filtration pressure and glomerular filtration rate are raised. When compared with the effect of an equivalent increase of effective filtration pressure by raising the arterial pressure, however, the effect of dilution is found to be six times as great in isolated kidneys and fifteen times as great in intact anæsthetised animals.‡ This appears to be due o an increased permeability of the glomerular membrane in the dilution liuresis, but the osmotic work is also increased, viz. about 40 per cent. when he flow of urine is doubled. In the intact dog there is also diminished reabsorption of water.

---

* Winton. *J. Physiol.*, 1934, **82**, 27P.

† Eggleton, Pappenheimer and Winton. *J. Physiol.*, 1940, **97**, 363.

‡ Eggleton, Pappenheimer and Winton. *J. Physiol.*, 1940, **98**, 336.

Water Diuresis. When a large amount of water is drunk, the change in he composition of the blood plasma may be very slight, and mainly due to oss of electrolytes diffusing out into the intestine and tissues, though the kidney is secreting large quantities of extremely dilute urine.* The diuresis is believed to be due to a reduction in the reabsorption of water by the tubules, rather than to an increase in glomerular filtration. Even when dilution of the plasma does occur, the maximal rate of urine formation lags some 20 to 30 minutes behind the maximal dilution,† so that the kidney responds slowly to the increased water content or lowered osmotic pressure of the blood.‡ It is possible to bring about a lowering of molecular concentration by injecting water or a hypotonic salt solution directly into the blood. Under these circumstances there is increased secretion of urine,§ but again only after considerable delay.‖

It is a curious fact that if the water reserves of the body are reduced by thirst, the administration of the amount of water lacking from the body produces a normal diuresis ¶ after the usual lag. It is clear, therefore, that the kidney does not respond directly to plasma changes, and the suggestion is that it is probably an indirect hormone effect.

Rôle of Pituitary. We have seen that the isolated perfused kidney secretes urine containing only traces of sodium chloride, although the content of this salt in the blood serum is rising steadily throughout the experiment. Addition of a small amount of extract of the posterior lobe of the pituitary body at once reduces the volume of the urine and increases the amount of chloride in it.** If, in such a perfusion, the blood is also circulated through an isolated head, the volume of urine is speedily reduced, and its chloride content correspondingly increased, but this does not occur if the pituitary body has first been removed from the head.

In the disease known as *diabetes insipidus* the patient passes large quantities of watery dilute urine and suffers very intense thirst as a consequence. Diabetes insipidus is often associated with demonstrable alterations, of a destructive nature, in the pituitary posterior lobe.†† Excision of the posterior lobe of the pituitary body brings about a similar condition, namely, increased formation of urine and disappearance of chlorides from the urine.‡‡ The action of the anti-diuretic pituitary hormone is to accelerate water reabsorption (in the distal tubule) and to retard sodium reabsorption (and hence also that of chloride) in the proximal tubule. Injection of posterior pituitary extract into a patient with diabetes insipidus,§§ suspends the condition, the urine diminishing in amount while the chlorides in it rise considerably and in some cases even exceed the percentage amount of chlorides in the plasma.

One milliunit of the pituitary anti-diuretic hormone is enough to stop diuresis due to water drinking : amounts of that order are believed to be usually present in the circulating blood,‖‖ and it is probable that water diuresis is due to suspension of liberation of the pituitary secretion into the blood, the delay in onset of diuresis being occasioned by the slow rate of destruction of the hormone.* It has been shown by Verney that the release of the anti-diuretic hormone from the pituitary into the blood is regulated in response to changes in the osmotic pressure of the blood plasma; if this is lowered (as by water drinking) the release is checked, and *vice versâ*. The receptors sensitive to osmotic changes in the blood (osmoreceptors) appear to be in the branches of the internal carotid artery † and injection of hypertonic solutions into the carotid

---

* Haldane and Priestley. *J. Physiol.*, 1916, **50**, 296; Rioch. *J. Physiol.*, 1930, **70**, 45.

† Smirk. *J. Physiol.*, 1933, **78**, 113, 127, 148.

‡ Baldes and Smirk. *J. Physiol.*, 1934, **82**, 62.

§ Fee. *J. Physiol.*, 1929, **68**, 305.

‖ Newton and Smirk. *J. Physiol.*, 1933, **78**, 451.

¶ Heller and Smirk. *J. Physiol.*, 1932, **76**, 1.

** Starling and Verney. *Proc. Roy. Soc.*, 1925, **97B**, 321.

†† Frank. *Berl. klin. Woch.*, 1912, **49**, 393.

‡‡ Vassale and Sacchi. *Riv. Sper. di Frenist*, 1892, **18**, 525.

§§ von den Velden. *Berl. klin. Woch.*, 1913, **50**, 2083.

‖‖ O'Connor and Verney. *Q. J. Exp. Physiol.*, 1942, **31**, 393.

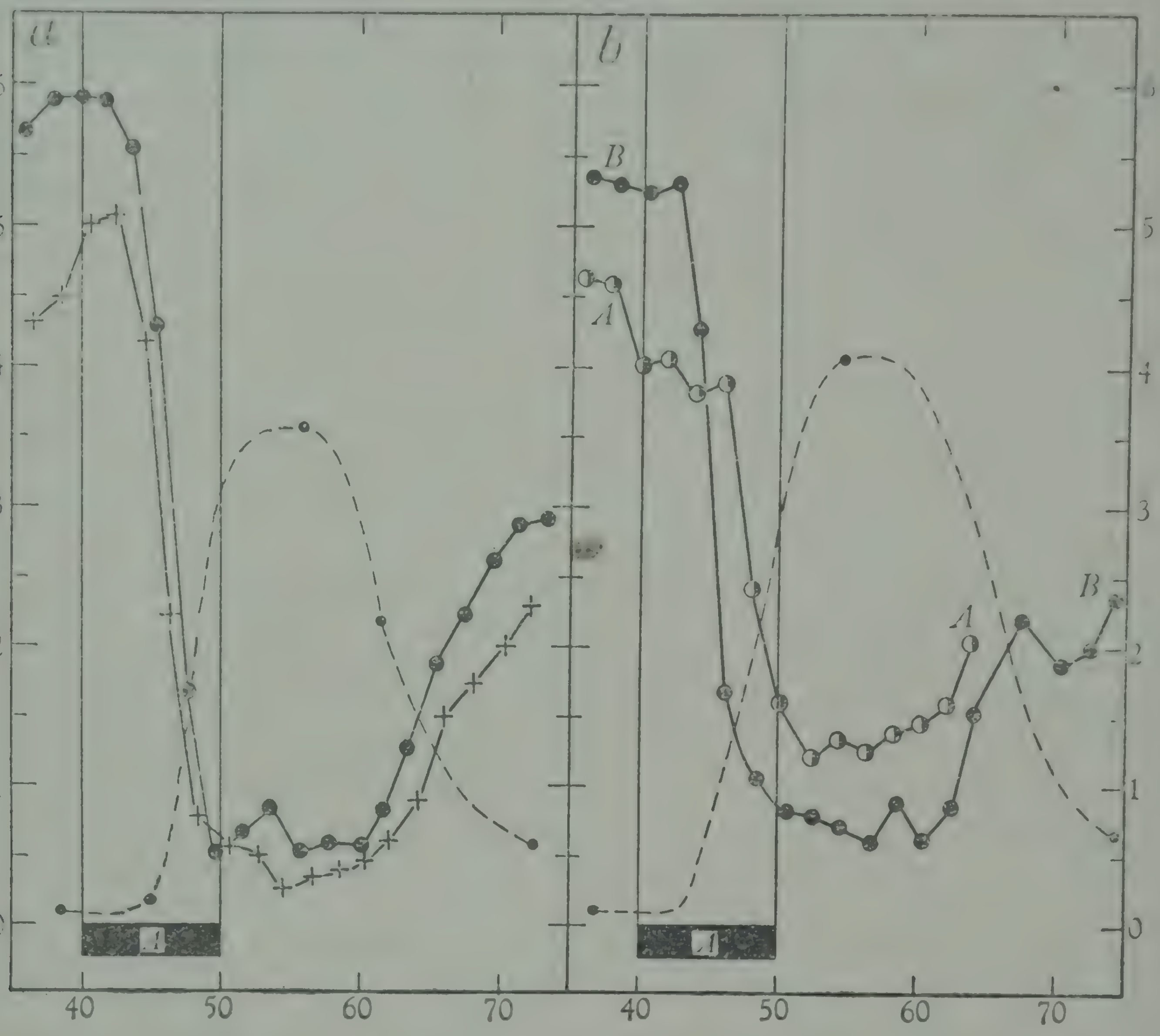


Fig. 618. Inhibition of water diuresis in dog.

(*a*) By injection into carotid artery, during the time A, of NaCl solution causing a 5 per cent. rise in osmotic pressure of the blood.

(*b*) by intravenous injection, during a similar period, of solution of post-pituitary extract, in curve A at rate of 1·67 μU/sec. and in curve B at 3·3 μU/sec. Broken line shows chloride concentration in urine. Abscissæ = time after the test dose of 350 c.c. water. Ordinates = urine flow.

(Verney. *Proc. Roy. Soc.*, 1947, **135B**, 85.)

artery soon cuts short a vigorous water diuresis. Water diuresis is also inhibited by emotional stress and this emotional inhibition is of two types, which together give a complex picture. One type of inhibition is again due to a reflex release of anti-diuretic hormone from the pituitary and is of slow onset and ending; it is abolished by a preceding injection of adrenaline, but not by sympathectomy. The other type is reflex, of quick origin and short duration and is abolished by denervation of the kidney or by splanchnic section.

Release of anti-diuretic hormone from the pituitary is controlled by impulses from the supraoptic nucleus of the hypothalamus. Injection of acetyl choline into this nucleus inhibits water diuresis in dogs with denervated kidneys, presumably by initiating such impulses.‡ These and similar facts remind us that the kidney is not an isolated organ, guided in its activity only through influences on the blood vessels, or by the presence of urinary constituents in the blood, but that it is played upon by hormones from distant parts of the body.

---

* Klisiecki *et al.* *Proc. Roy. Soc.*, 1933, **112B**, 496.
† Verney. *Proc. Roy. Soc.*, 1947, **135B**, 25.
‡ Pickford. *J. Physiol.*, 1947, **106**, 264.

Action of Diuretics. Attempts have been made to illuminate the problem of renal secretion by studying the action of diuretics, *i.e.* substances which, injected into the blood stream or absorbed from the alimentary canal, increase the secretion of urine. In anæsthetised animals, a large increase in the urinary flow can be brought about by the intravenous injection of crystalloids such as sodium sulphate, potassium nitrate, or sugar. When a concentrated solution of salt is injected into the general circulation, water is drawn from the tissues, *i.e.* a condition of hydræmic plethora is set up. The kidney placed in an oncometer shows a great increase in volume. While the plethora lasts, there are mechanical conditions at work in the kidneys, *i.e.* rise of pressure, greater rate of blood flow, and diminished concentration of plasma—which would concur in producing an increased glomerular transudation. With certain salts, such as sodium chloride, the diuresis may be coterminous with the hydræmic plethora, but with other members of this class, such as glucose, the diuresis always outlasts the plethora, so that the continued augmentation in the secretion of urine leads to an actual concentration and diminution of the volume of the circulating blood. This behaviour in anæsthetised animals is due to the pituitary being out of action. In man, crystalloid solutions do not cause much diuresis.

As we have already seen, every substance which is not absorbed from the glomerular filtrate by the tubules should act as a diuretic since it will oppose osmotic resistance to the absorbing powers of the cells. Thus the low-threshold substances, urea, and sodium sulphate, nitrate, and phosphate, will, in any concentration, act as diuretics. The threshold substances will produce diuresis so long as their concentration in the plasma surpasses their normal threshold value. In urea diuresis, the glomerular filtration is almost unaltered, as is shown by the creatinine clearance ; the osmotic work is increased, but the oxygen consumption is not, *i.e.* the efficiency is increased. For a doubling of urine flow in the isolated kidney the osmotic work is increased 65 per cent. and the efficiency 60 per cent.* by urea diuresis.

Hayman and Starr † have counted in the rabbit's kidney the number of glomeruli which were open to circulation when in the normal state, and also after administration of caffeine. They find that, whereas in the normal state, from 50 to 90 per cent. of the glomeruli are patent, all of them are opened after a dose of caffeine. After adrenaline, on the other hand, only 10 per cent. of them may be in use, and the urine flow much reduced.

On the other hand, it seems certain that most diuretics—including caffeine—act on the tubules, causing a paralysis of their powers of absorption. Phlorrhizin may be assumed to paralyse the absorptive powers of the tubular cells for glucose. The glucose in the glomerular transudate, in passing through the tubules, may thus be concentrated sixty to a hundred times. We may assume that different diuretics have similar powers of paralysis on the absorptive mechanisms of the tubules, either general, or confined, as in the case of phlorrhizin, to one or other of the normal constituents of the plasma.

Renin. Saline extracts of alcohol-dried kidneys contain a substance which raises arterial blood pressure and accelerates urine formation owing to reduction of tubular reabsorption of water and sodium chloride.‡ The active substance, called *renin*, is an enzyme, and it has been suggested that it plays a part in the production of the rise of blood pressure, with diuresis, which is a feature of certain kidney diseases. Reduction of the blood flow to a kidney, by clamping, leads to a state of hypertension similar to the pathological condition, and is attributed to release of renin or some related substance into the blood stream. The release, into the blood, of a vasoconstrictor substance by isolated perfused kidneys has been demonstrated.§ (*v.* also p. 659.)

---

* Eggleton, Pappenheimer and Winton. *J. Physiol.*, 1940, **97**, 363.

† Hayman and Starr. *J. Exp. Med.*, 1925, **42**, 641.

‡ Pickering and Prinzmetal. *J. Physiol.*, 1940, **98**, 314.

§ Verney and Vogt. *Q. J. Exp. Physiol.*, 1938, **28**, 253.

For Reference

Blalock (1940). Experimental Hypertension. *Physiol. Rev.*, **20**, 159.

Cushny (1926). "The Secretion of Urine." (London: Longmans Green.).

Goldblatt (1947). The Renal Origin of Hypertension. *Physiol. Rev.*, **27**, 120.

Marshall (1926). The Secretion of Urine. *Physiol. Rev.*, **6**, 440.

Marshall (1934). The Comparative Physiology of the Kidney, etc. *Physiol. Rev.*, **14**, 133.

McCance (1948). Renal Function in Early Life. *Physiol. Rev.*, **28**, 331.

Herrin (1941). Factors affecting the Tests of Kidney Function. *Physiol. Rev.*, **21**, 529.

Homer Smith (1937). "The Physiology of the Kidney." (New York and London: Oxford University Press.).

Homer Smith (1940). The Physiology of the Renal Circulation. *Harvey Lects.*, p. 166. (Lancaster, Pa. Science Press Printing Co.).

Homer Smith (1943). Lectures on the Kidney. (Univ. Kansas, Univ. Exten. Div.).

O'Connor (1947). The Control of Urine Secretion in Mammals by the Pars Nervosa of the Pituitary. *Biol. Rev.*, **22**, 30.

Pickford (1945). Control of the Secretion of Antidiuretic Hormone from the Pars Nervosa of the Pituitary Gland. *Physiol. Rev.*, **25**, 573.

Pitts (1948). Renal Excretion of Acid. *Feder. Proc.*, **7**, 418.

Richards (1938). *Croonian Lect., Royal Society, Proc. R. S.*, 1938, **126B**, 398.

Shannon (1939). Renal Tubular Excretion. *Physiol. Rev.*, **19**, 63.

Winton (1937). Physical Factors involved in the Activities of the Mammalian Kidney. *Physiol. Rev.*, **17**, 408.

Yuile (1942). Hemoglobinuria. *Physiol. Rev.*, **22**, 19.

3. THE PHYSIOLOGY OF MICTURITION

The urine as it is formed passes through the ureters to the bladder, where it gradually accumulates, and is voided at intervals.

The Ureters enter the bladder at its base, at the two posterior angles of the trigonum. Their entrance is by an oblique valvular orifice which prevents reflux of urine from bladder to ureter. The plain muscular coats of the ureter show rhythmic contraction waves; these pass from the pelvis of the kidney to the bladder at a velocity of about 2 to 3 cms. per second, driving the contained fluid in front of them, so that the pelvis of the kidney is emptied of fluid. The intermittent injection of urine into the bladder as each wave of contraction reaches it can be seen on cystoscopic examination in man. The frequency of the contractions, normally about three per minute, is increased by warming the ureter, and by distension, so that the waves are more frequent when the secretion of urine is profuse. The ureters are supplied with nerve fibres from the splanchnic nerves by way of the renal plexus, and at their lower ends from the hypogastric nerves. Stimulation of the hypogastrics closes the openings into the bladder.

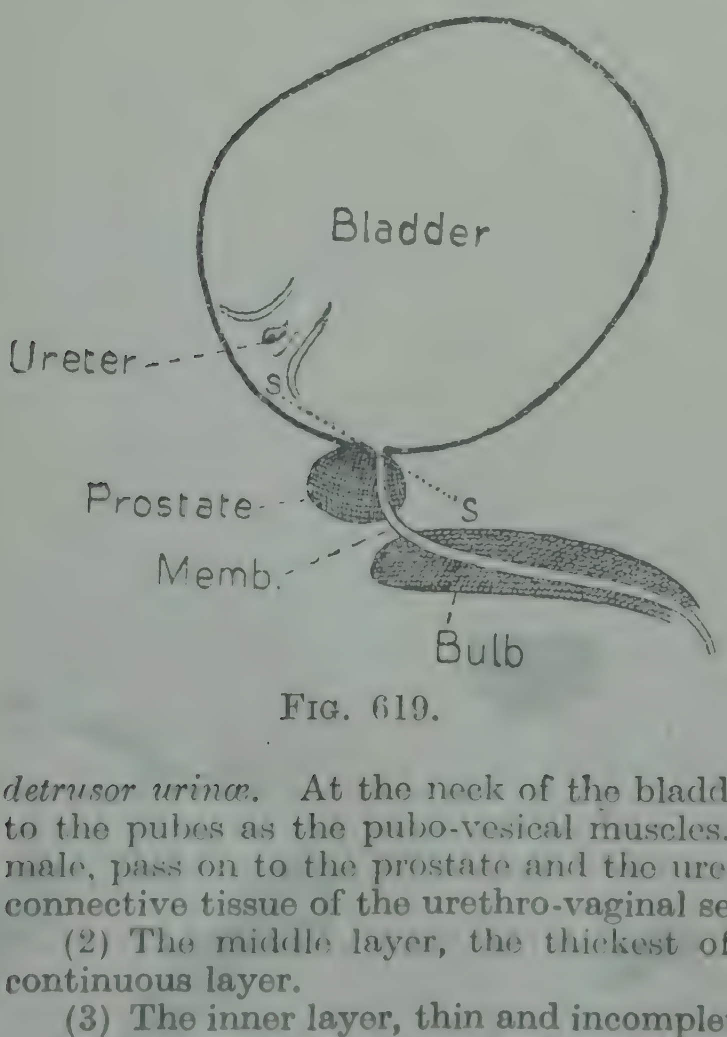


Fig. 619.

The Bladder is lined by transitional epithelium, closely adherent to the underlying muscular coat. In the latter are three layers of muscular fibres:

(1) An outer layer, running longitudinally from the neck of the bladder to the fundus, sometimes distinguished by the name of the *detrusor urinæ*. At the neck of the bladder these bundles send some fibres to be attached to the pubes as the pubo-vesical muscles. On the dorsal surface some bundles, in the male, pass on to the prostate and the urethra, while in the female they end in the tough connective tissue of the urethro-vaginal septum.

(2) The middle layer, the thickest of the three, arranged circularly and forming a continuous layer.

(3) The inner layer, thin and incomplete.

The name of *detrusor* ought physiologically to be applied to the whole of the three coats, which act as one in diminishing the capacity of the bladder.

THE SPHINCTERS. At the base of the bladder, the middle layer is modified over the trigonum by fibres, which serve as the most important sphincter, the *sphincter vesicæ*, or *internal sphincter*. Its fibres pass obliquely forwards and downwards and form a loop around the orifice of the bladder, lying on the ventral side of the bladder below and quite distinct from the thick coat of circular fibres belonging to the bladder itself (SS, Fig. 619; Fig. 620). If a catheter be passed into the urethra no urine escapes until its tip has passed the internal sphincter and actually entered the bladder. The sphincter muscle in the male is continuous with the muscular tissue of the prostate, and in the female with the muscle of the urethral wall; these continuations of the sphincter, by their tonic contraction, will also tend to prevent the escape of urine along the canal.

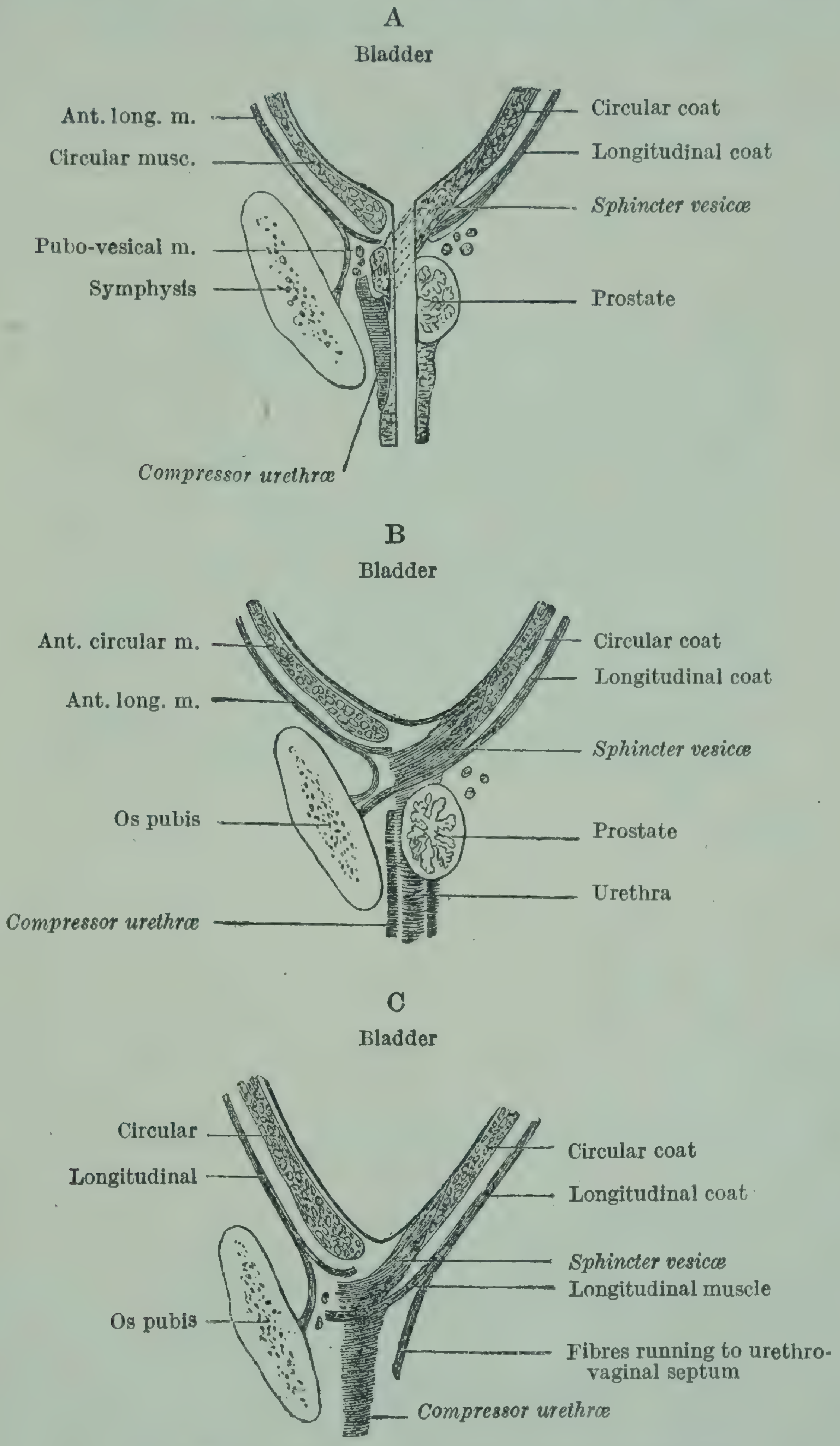

A
Bladder
Ant. long. m.
Circular musc.
Pubo-vesical m.
Symphysis
Compressor urethræ
Circular coat
Longitudinal coat
Sphincter vesicæ
Prostate
B
Bladder
Ant. circular m.
Ant. long. m.
Os pubis
Compressor urethræ
Circular coat
Longitudinal coat
Sphincter vesicæ
Prostate
Urethra
C
Bladder
Circular
Longitudinal
Os pubis
Circular coat
Longitudinal coat
Sphincter vesicæ
Longitudinal muscle
Fibres running to urethro-vaginal septum
Compressor urethræ

FIG. 620. Saggital Sections through Neck of Bladder. (METZNER after KALISCHER.)

A. In middle line (male). B. Slightly to left of middle line (male). C. Ditto (female).

This urethral muscle is strengthened by two other sphincter muscles, which are composed of striated fibres. Together they form the *external sphincter*. They can be closed voluntarily, but are only opened reflexly and as part of the act of micturition. The chief one, the *compressor urethræ*, forms a flat ring around the second part of the urethra, extending in the male from the prostate to the bulb, where its function is taken up by the *bulbo-cavernosus* (Fig. 620).

The bladder is therefore supplied with a powerful muscular wall, the contraction o which will cause its evacuation, and with sphincters of two kinds, one involuntary, th internal *sphincter*, at the neck of the bladder, and two voluntary external sphincters, th *compressor urethræ* and *bulbo-cavernosus* muscles, which can empty the lower parts of th urethra. In man, after destruction of the first part of the urethra by suprapubic prostatectomy, the compressor urethræ is the competent sphincter.

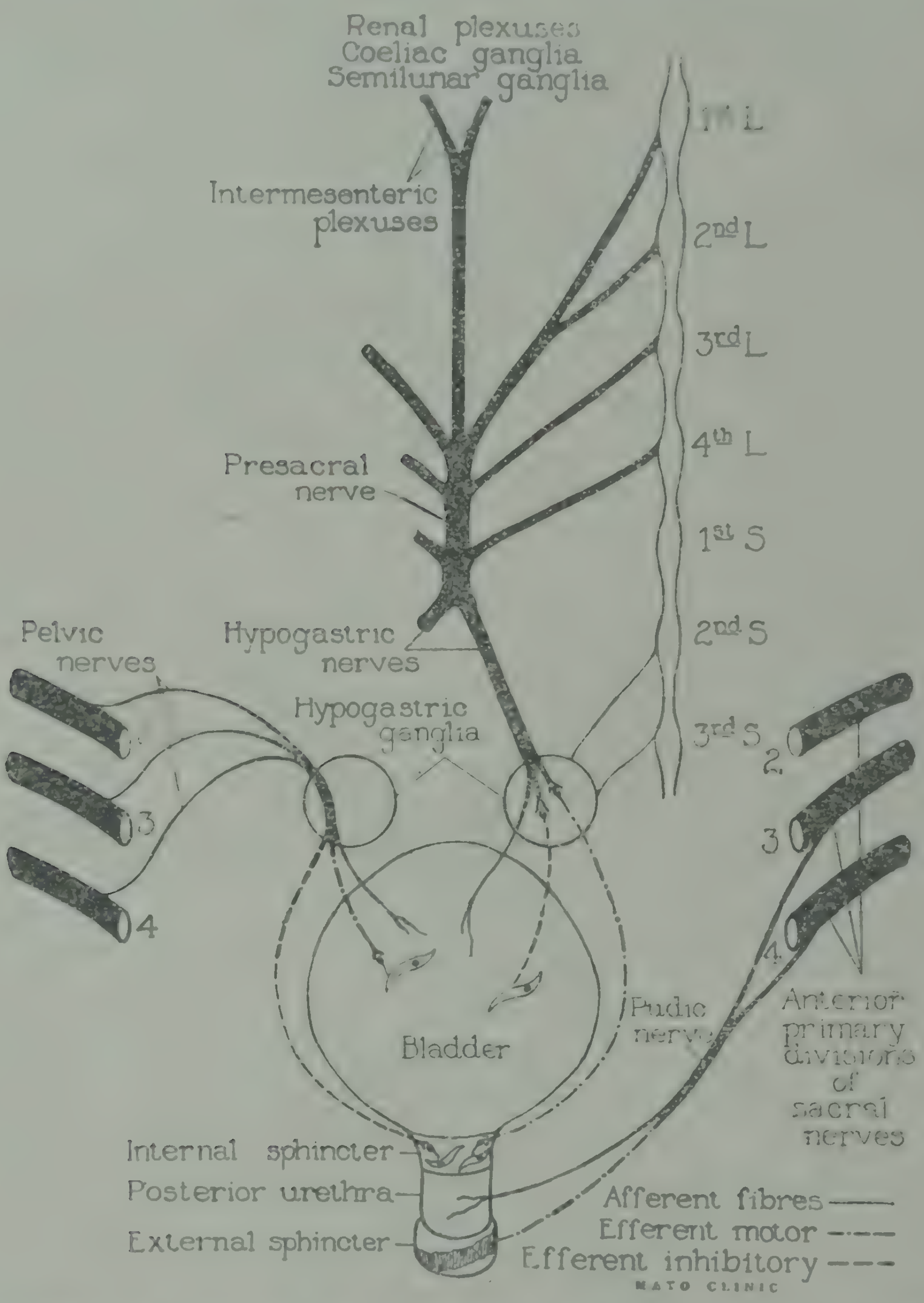

Renal plexuses
Coeliac ganglia
Semilunar ganglia
Intermesenteric plexuses
1st L
2nd L
3rd L
4th L
1st S
2nd S
3rd S
Presacral nerve
Pelvic nerves
Hypogastric nerves
Hypogastric ganglia
2
3
4
Bladder
Pudic nerve
Anterior primary divisions of sacral nerves
Internal sphincter
Posterior urethra
External sphincter
Afferent fibres
Efferent motor
Efferent inhibitory
MAYO CLINIC

FIG. 621. Diagram of Innervation of the Bladder and Sphincters in Man. (After LEARMONTH.)

The Nerve Supply of the Bladder and internal sphincter (Fig. 621) is derived from two main sources, namely, from the hypogastrics (sympathetic) and from the pelvic visceral nerves or *nervi erigentes*. The upper two or three lumbar nerves send white rami communicantes to the lateral chain of the sympathetic, and thence to the collateral plexus of sympathetic nerves and ganglia which is connected above with the aortic plexus and which below forms a ganglionic mass often referred to as " the intermesenteric plexus." Most of the fibres then pass, by the plexus of hypogastric nerves, into the pelvis on each side of the rectum and in man end by relaying in the hypogastric ganglia, at the base of the bladder. From thence, fibres pass to the bladder wall. In the cat the hypogastric nerves are post ganglionic, with cell stations in the inferior mesenteric ganglia.

The pelvic visceral nerves are derived from the second and third sacral nerves. They pass directly to the vesical plexus and are carried with branches of this plexus to the neck of the bladder. The fibres do not run directly from the spinal cord to their ending in the bladder wall, but are interrupted by cells situated peripherally, partly in the vesical plexus, but chiefly in the walls of the bladder itself. Both sets of fibres carry efferent impulses to the bladder, and also supply the rectum and the colon. Afferent impulses from the bladder and urethra travel in both, but chiefly in the pelvic visceral nerves, and enter the spinal cord by the dorsal roots of the sacral nerves, particularly the second and third.

The afferent and efferent nerves of the urethra travel in the pudic nerves and in the hypogastrics.

The muscles of the external sphincter are supplied by the pudic nerve.

Action of the Pelvic Visceral Nerves. About a third of the fibres of the pelvic nerve are afferent. Excitation of the peripheral end of one pelvic visceral nerve causes a strong contraction of the same side of the bladder, sometimes extending also to a slight extent to the contralateral half. When both pelvic nerves are stimulated simultaneously, contraction of both sides of the bladder causes a rise of pressure in its interior (Fig. 622) which is sufficient to overcome the resistance of the sphincter and to cause a complete emptying of the bladder.

Relaxation of the sphincter is also believed to be brought about by impulses descending the pelvic visceral nerves. There is no doubt, therefore, that these nerves are most important for the act of micturition.

Section of the nerve on one side causes no abnormality in micturition. Section of both pelvic nerves paralyses micturition, but power of partial evacuation of the bladder may return in a few weeks. If the hypogastrics be now cut, or even the sacral cord extirpated, the bladder is not completely paralysed, though its evacuation becomes unconscious and incomplete.

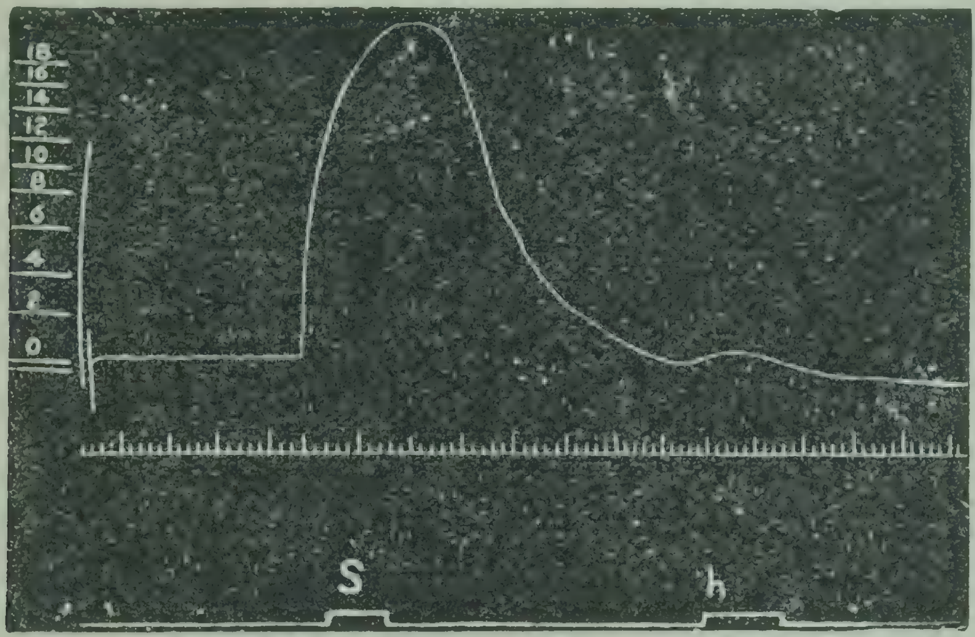


Fig. 622. Curve showing Rise of Pressure in the Bladder caused by Stimulation of S, Pelvic Visceral Nerves ; *h*, Hypogastric Nerves. (Fagge.)
The scale indicates centimetres of water.

Action of the Hypogastric Nerves. These nerves show differences in their action according to the animal and the conditions. In nearly all animals, including man, stimulation of the hypogastric nerves causes contraction of the trigone, ureteric openings, and internal sphincter. But in man, and many other animals, it may also excite either contraction or relaxation (or both) of the detrusor, and so cause either a rise * or a fall of pressure in the bladder. Complete removal of the sympathetic supply to the bladder in man appears at first to result in some weakness of the internal sphincter, but this soon passes off and normal micturition is restored. Further, since stimulation may relax the detrusor in some cases, the retention of urine and the distension of the bladder may be aided by a reflex relaxation of the bladder wall and a reflex constriction of the sphincter, excited in each case through these nerves. We may therefore regard these nerves as normally and chiefly exercising a tonic effect on the internal sphincter, and micturition must be associated with a central inhibition of this tonic activity.

Normally, therefore, both sets of nerves are essential. The hypogastrics, by their undoubted action on the internal sphincter, play an especially active part during the accumulation of urine in the bladder, while the pelvic visceral nerves are necessary for the complete evacuation of the bladder which occurs at micturition.

Sensory fibres run in both nerves : those for pain of stretching of the bladder in the sympathetic, and those for touch, together with afferents involved in the micturition reflex, in the pelvics.

---

* Mellanby and Pratt. *Proc. R. S.*, 1939, **127B**, 307 ; 1940, **128B**, 186.

The Contraction of the Bladder. The muscle of the bladder wall, being plain muscle, executes slow contractions. The nature and power of these vary according to whether the contractions are "isotonic" (constant pressure in bladder) or "isometric" (bladder contents kept at constant volume).*

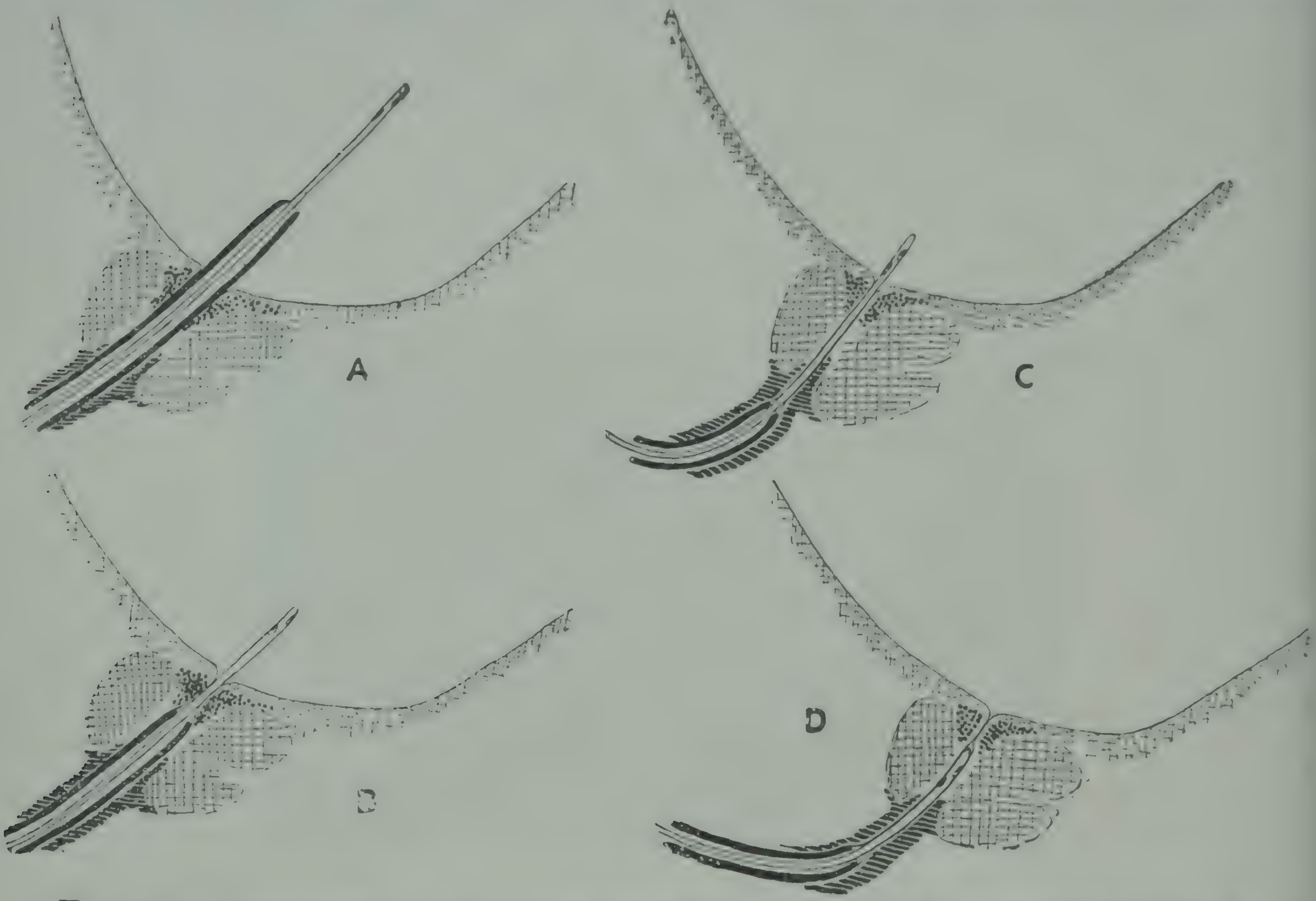


FIG. 623. Four arrangements of concentric double catheters for studying pressures in bladder, action of sphincters, etc. Internal sphincters stippled, external sphincter heavily hatched, prostate lightly cross-hatched. The wide catheter *a* is connected with the tube *f* (Fig. 624), and the narrow catheter *b* running inside it, with the bulb *l*. Fig. A: both catheters in bladder, one for filling or emptying, one for recording pressure. Fig. B: only internal sphincter functional. If a leak is arranged in the outlet of the larger catheter, a rise of pressure in that catheter shows that the internal sphincter is relaxing and allowing fluid to leave the bladder. Similarly in position C and D, the relaxation of the external sphincter can be demonstrated. (DENNY-BROWN and ROBERTSON, *Brain*, 1933, **56**.)

Under isotonic conditions, the normally innervated bladder exhibits regular rhythmic variations of volume, lasting some ten seconds, which increase in frequency and diminish in amplitude when the filling pressure is raised. Stimulation of the hypogastrics, or adrenaline in large doses, causes relaxation; stimulation of the pelvic nerves, or acetylcholine, causes contraction (but if the pressure maintained in the bladder is high, acetylcholine may cause relaxation).

Under isometric conditions, the bladder often remains quiescent for long periods, which alternate with powerful contractions, raising the pressure often to more than 100 mm. Hg, and lasting for a minute or more. These are abolished by section of the pelvic nerves and are probably reflex. Changing over from the isotonic to the isometric state at the moment when an isotonic contraction is developing leads at once to a powerful isometric contraction, but if the change-over is made when an isotonic contraction is subsiding, a period of isometric rest at zero pressure commences.

It is probable that the initiation of powerful isometric contractions whenever the pressure in the bladder rises suddenly is a reflex from stretch receptors in

---

* J. P. EVANS. *J. Physiol.*, 1936, **86**, 396.

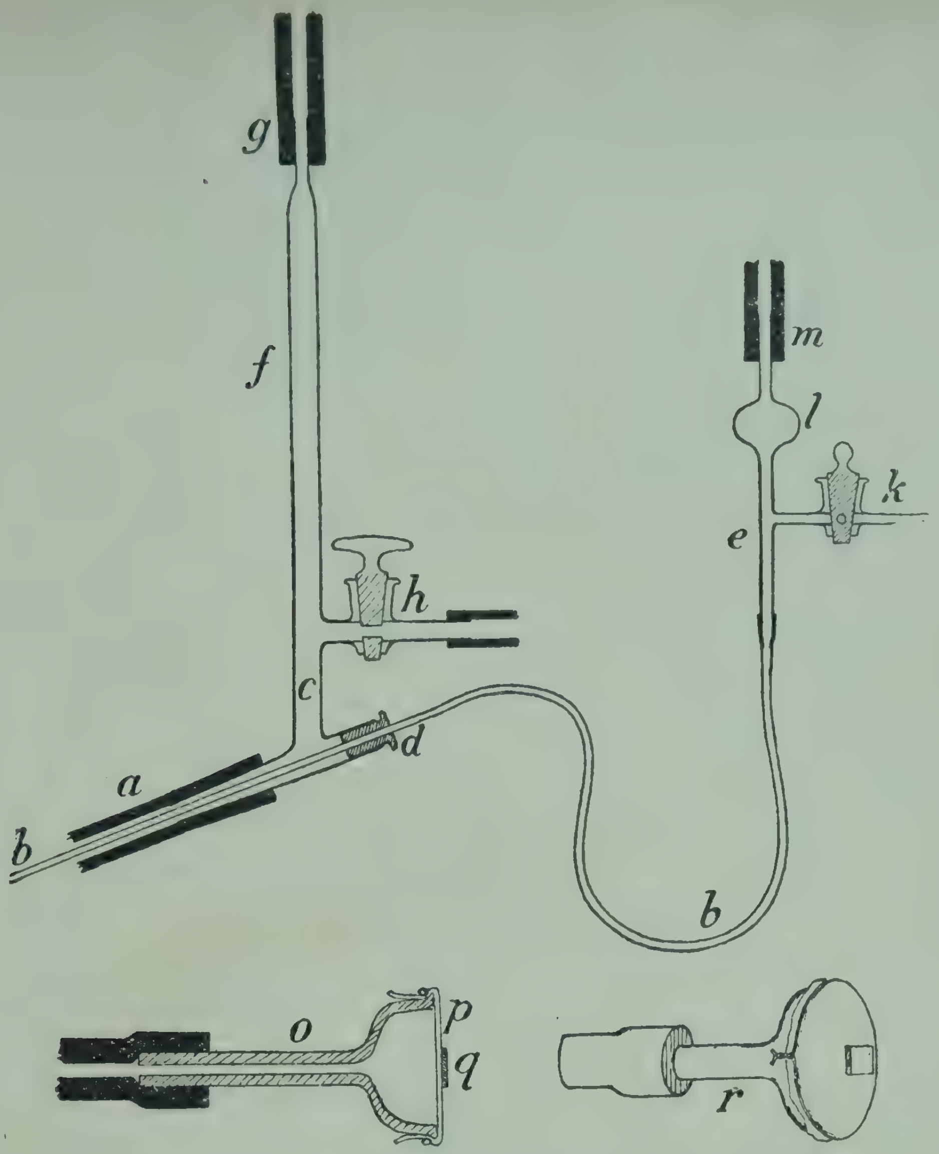


Fig. 624 (*Top*). Filling and recording system for human bladder. Reservoirs of fluid connected and controlled through taps *h* and *k*.

(*Below*.) Optical manometers connected with *g* and *m*. *p*, rubber membrane; *q*, mirror. (Denny-Brown and Robertson.)

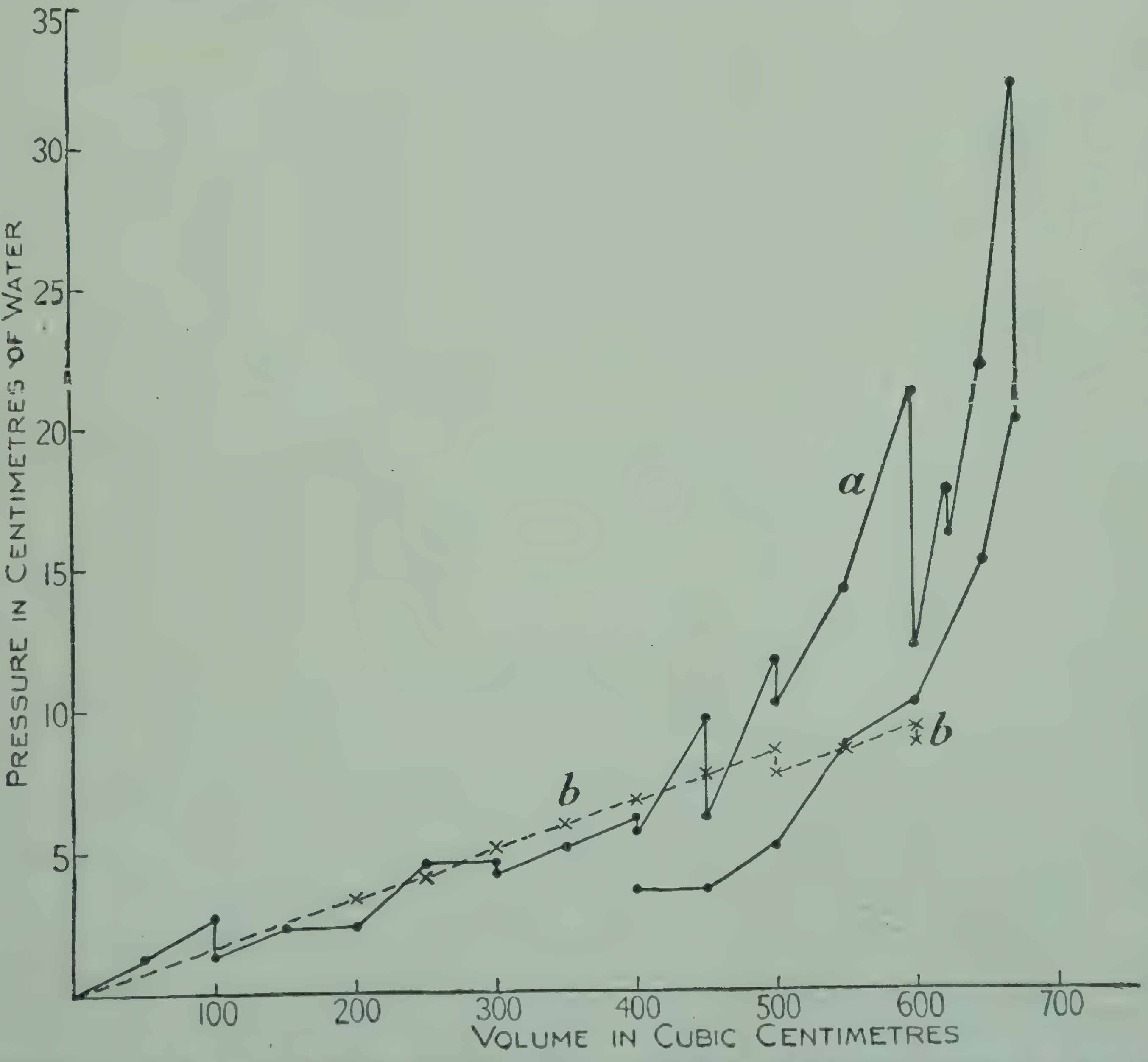


Fig. 625. *a*, Record of filling, followed by emptying, of normal bladder, in 50 c.c. steps. The vertical drops indicate adaptation during the intervals. *b*, Record of a filling just after bladder had been subjected to prolonged preliminary distension. (Denny-Brown and Robertson.)

the bladder wall. Afferent impulses along the pelvic nerves have been demonstrated when the bladder pressure is raised,* and once the reflex is started it is evidently self-exciting, and probably plays an important part in micturition.

The Filling of the Bladder. Except during micturition, both sphincters are in a state of tonic contraction. Thus, it requires a considerably greater pressure in the bladder to overcome the resistance of the sphincters during life than after death. The urine, therefore, as it is secreted will normally accumulate in, and distend, the bladder. Let us first consider the filling of the normal bladder. The filling and emptying of the bladder in man can be studied by the introduction of appropriate catheters as shown in Fig. 623, the catheters being connected with optical recording manometers, shown in Fig. 624. The normal bladder with its proper innervation accommodates a good deal of fluid before showing any considerable rise of pressure, or pronounced contractions. Even when fluid is introduced by a catheter 50 or 100 c.c. at a time, there is only a momentary tonic resistance, which soon subsides (Fig. 625). With a filling of about 400 c.c. the pressure rises more sharply, and at 600 to 800 c.c. there are powerful rhythmic contractions accompanied by painful sensations. By voluntary effort these can be suppressed for a time (Fig. 626), but when the pressure reaches about 100 cm. water micturition follows involuntarily.

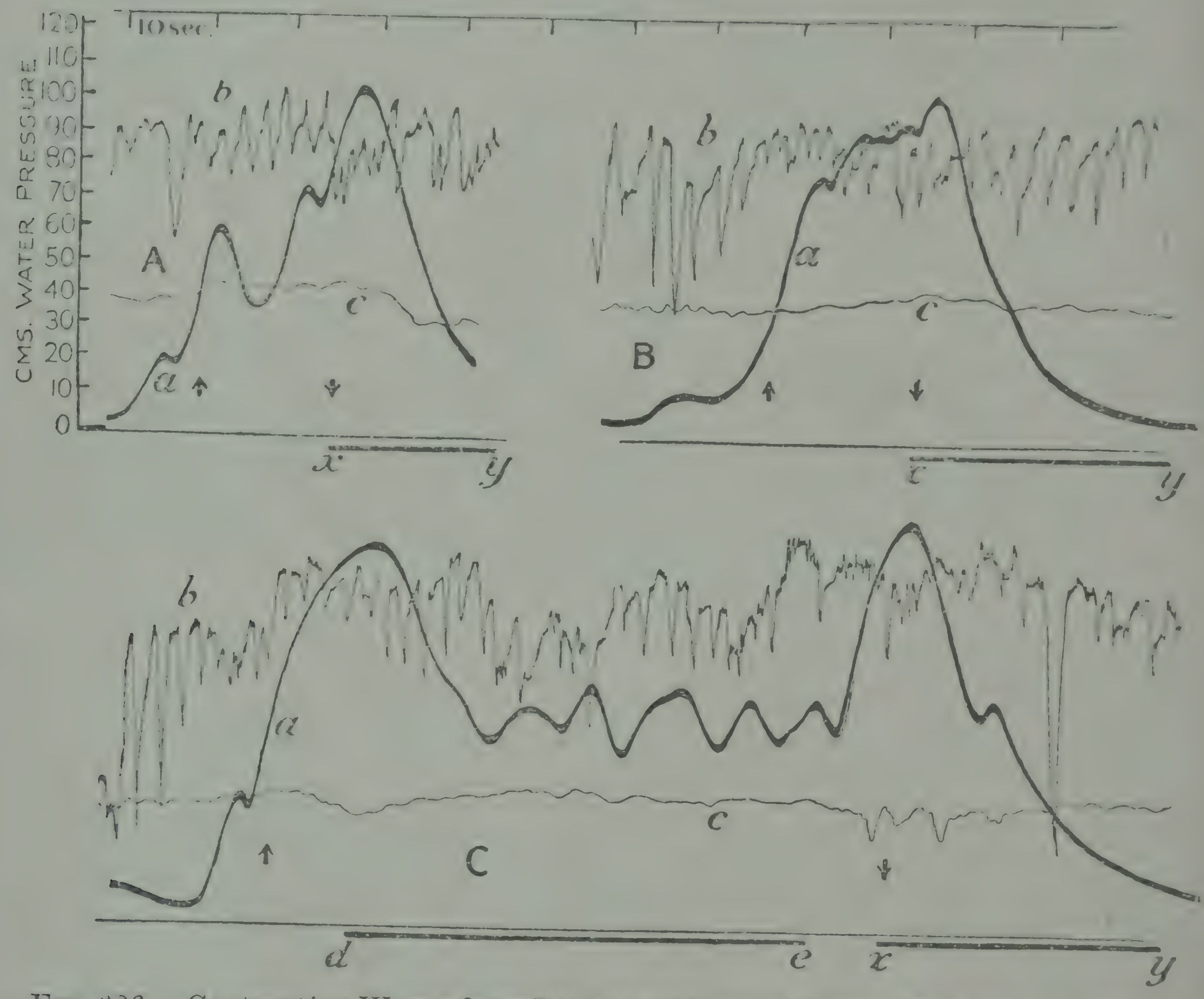


FIG. 626. Contraction Waves from Bladder of Normal Human Subject. *a*, vesical pressure (catheter in bladder); *b*, abdominal movements; *c*, rectal pressure. Arrows pointing up and down indicate opening and closing respectively of sphincters. $x - y$ = strong and $d - ɜ$ weaker, efforts at restraint. (Note reflex bladder contractions when sphincter is relaxed.) (DENNY-BROWN and ROBERTSON.)

The filling conditions are different in a bladder cut off from the central nervous system, as in patients some time after a lesion of the cauda equina and sacral cord. We may then regard the bladder as a tonic viscus with elastic walls containing plain muscle, and with a tonically contracted sphincter. If a distending force be applied rapidly, the sudden increase of tension acts as a direct excitant to the muscle, causing it to enter into contraction, which may be tonic or rhythmic. The effect of distension of the denervated bladder on the tension in this organ will depend, therefore, on the rapidity with which it occurs. In any case, the denervated bladder is highly tonic and shows a considerable rise of pressure when fluid is injected into it by a catheter; thus 200 c.c. will give a pressure of 70 cm. water, which later may fall by slow adaptation to 30 cm. or less. At such pressures the effect of the stretching of the muscular wall makes its appearance. A manometer connected with the bladder shows a series of rhythmic contractions of the muscular wall (Fig. 627, A), each lasting about a minute, and becoming more powerful as the distension of the bladder augments (Fig. 627, B). In the bladder entirely cut off from its connection with the central nervous system, these automatic rhythmic contractions gradually increase in force, until one of them suffices to overcome the resistance presented by the tonically contracted sphincter. A partial emptying of the bladder therefore takes place, but before the bladder has been emptied, the pressure falls below that necessary to overcome the resistance of the sphincter, so that there is always, under these circumstances, a certain amount of residual urine left in the bladder. This is the condition found in animals when the lower part of the spinal cord has been extirpated, or in man when it has been destroyed as the result of accident or disease.

---

* MCCREA and MACDONALD. *J. Physiol.*, 1936, **86**, 20P; LEARMONTH. *Brain*, 1931, **54**, 147.

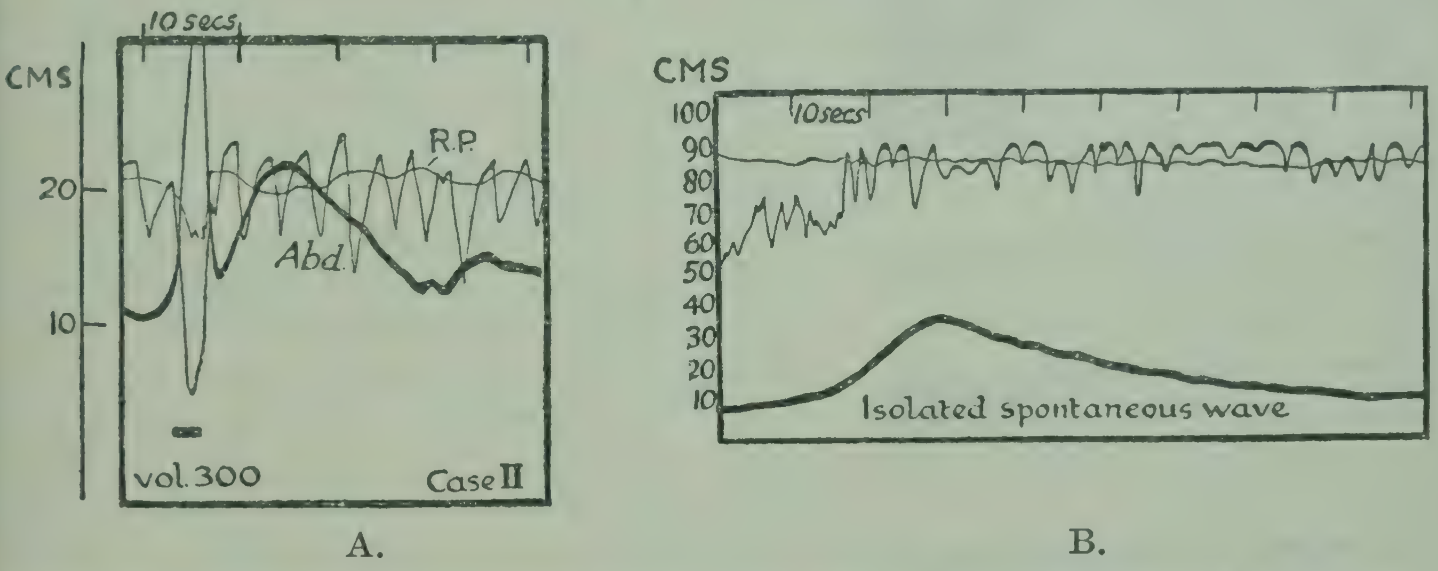


Fig. 627. Bladder Contractions from Patients with Lesions of the Cauda Equina. A, Simplest type of contraction: elicited by brief passive abdominal pressure. Duration of stimulus shown by short horizontal line. B, Lowest grade of compound wave: spontaneous. R.P., Rectal pressure. Abd., Abdominal movements. (Denny-Brown and Robertson.)

Micturition. In the denervated bladder, the factor finally causing partial evacuation is the gradual increase in the intravesical tension from the accumulation of fluid. The same factor is prepotent in determining the onset of micturition at lower pressures in the normal animal. The desire to micturate generally arises in normal adults when about 400 c.c. of urine has collected in the bladder; it can be voluntarily overruled until the amount is much greater, but at about 700 c.c. becomes urgent and painful. The main path for these sensations is by the sympathetic, and the act of micturition is a reflex which is liberated from voluntary suppression. As a preliminary to this release the perineal muscles are relaxed. In the absence of restraint from the higher centres (*e.g.* in the infant) micturition will take place each time that the tension in the bladder has reached a certain height, the amount of fluid in the bladder at the time depending on the one hand, on the rate at which the fluid has entered this organ from the ureters, on the other hand, on the irritability of the bladder wall itself and of the nervous centres concerned with its motor innervation. The effect of the gradual accumulation of fluid and rise of tension i twofold. In the first place, it acts on the bladder wall, causing rhythmi contractions of ever-increasing intensity; in the second place, the mere stretching of the bladder originates impulses in the sensory nerve-endings in its wal and these, as we have seen with the "isometric" bladder, are reinforced a every rise of tension caused by the rhythmic contractions. These impulse travel up to the nervous centres, and are summated until, when released, they result in a sudden discharge of efferent impulses of two kinds, namely:

(1) Motor impulses to the whole musculature of the fundus of the bladder.

(2) Inhibition of the tonic contraction of the sphincter muscles. The resultant of these two processes, the contraction of the detrusor and the relaxation of the sphincters, is a complete emptying of the bladder, and the act is completed by the contraction of the involuntary and voluntary muscles surrounding the urethra and causing complete expulsion of the contents of this tube.

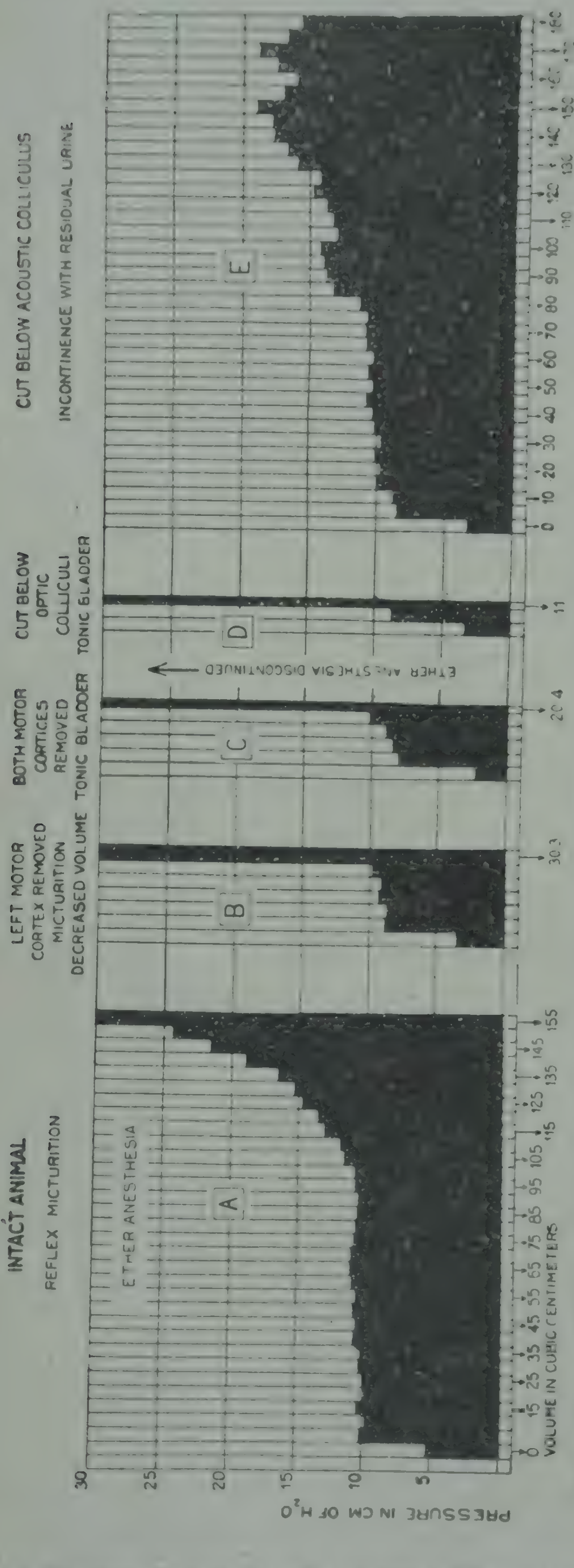


FIG. 628. Effect of brain sections on volume-pressure relations and emptying of bladder in etherised cat. (LANGWORTHY and KOLB, *Brain*, 1933, **56**, 731.)

Observations on man show that an active relaxation of the internal sphincter is a necessary part of the act of micturition. Thus a rigid catheter was introduced into the bladder, which was fully distended with fluid. On withdrawing the catheter until its opening lay just outside the internal sphincter, the flow of urine stopped. The man, however, was able to micturate directly he was told, and to stop again, though less suddenly, at will. It was impossible in this case for any of the urethral muscles to be concerned, since the rigid catheter impeded their action.

The Reflexes of Micturition. The investigations of Barrington * have revealed six closely co-ordinated reflexes concerned in micturition in the cat, viz.:—

(1) *Contraction of the bladder* when its internal pressure approaches 10 cm. of water. The afferent and efferent paths are in the pelvic nerves.

(2) *Contraction of the bladder*, even when only partially full, as a result of running fluid through the urethra. This ensures that when micturition is started by the first reflex, it will normally be continued to completion. The afferent path is the pudic, the efferent the pelvic nerve.

(3) *Contraction of the bladder* (slight), when the first part of the urethra is distended. Both paths are in the hypogastrics.

---

* BARRINGTON. *Q. J. Exp. Phys.*, 1914, **8**, 33; 1915, **9**, 261; 1925, **15**, 81; *Brain*, 1921, **44**, 23; 1922, **45**, 126; 1928, **51**, 209; 1931, **54**, 177.

(4) *Relaxation of the urethra* when fluid passes along it. Both afferent and efferent paths are in the pudic nerves.

(5) *Relaxation of the urethra* when the bladder contracts. The afferent path is in the pelvic nerves, the efferent in the pudic.

(6) *Relaxation of plain muscle of proximal third of the urethra*, when the bladder is distended. Both paths are in the pelvic nerves.

It will be seen that each of these reflexes, except the third, is capable of giving rise to the remainder, and so collaborating to ensure normal micturition. There is little doubt that similar reflexes are operative in man.

THE CENTRAL CONTROL OF THE BLADDER. Like many complex reflexes, micturition is under the joint control of higher and subsidiary nerve centres; in all there are at least four levels of control—namely, cortical, hypothalamic, brain-stem and spinal.

The nature of this control is illustrated in Fig. 628, which shows the progressive rise of bladder pressure in the cat as the contents are incremented in 5 c.c. steps, until finally evacuation occurs, complete or otherwise. With the intact animal the reflex occurred at 155 c.c.; after removal of one or both cerebral cortices the reflex was much more easily started and relaxation to accommodate the contents was slight. Section below the superior colliculi, led to so excitable a reflex that evacuation took place at only 11 c.c., but section below the inferior colliculi gave an inert bladder, showing retention with overflow, *i.e.* with incomplete emptying and no proper reflex.

In the human adult, the processes of retention and evacuation of urine are modified and controlled by reflexes and by voluntary effort. The human bladder shows a readier accommodation to increase of its contents without undue rise of pressure when its nerve supply is intact than when it has been denervated, even though connections with higher centres have been cut off as, for instance, after section of the cord in the mid-dorsal region. But its capacity is still further improved by the influence of the higher centres, which can exercise considerable restraint over the reflex micturition mediated by the lower centres. The normal action of the sphincter mechanism may be aided by the contraction of the perineal muscles, which keep the urethra closed. Under usual conditions, however, the intact subject micturates before much conscious restraint has had to be exercised. In normal micturition there is first of all a voluntary removal of restraint, so releasing the reflexes of micturition which are effected by the lower centres. These reflexes involve the contraction of the bladder and the release of the sphincters. If the contents of the bladder are small in volume, the reflex process of evacuation may be set in motion by voluntary contraction of the abdominal muscles, by which the pressure in the bladder is increased and the sphincter action reflexly inhibited (reflexes 5 and 6 above).

The third, fourth, fifth and sixth of the reflexes described by Barrington are spinal, the centre concerned being situated in the lumbo-sacral spinal cord. If this centre and its connections be intact, reflex micturition may be carried out even after section of the cord in the dorsal region, but according to Barrington, the process of micturition is always imperfect in animals and in man after section of the cord in any part of its course. In Barrington's first and second reflexes the dominant centre in the cat is in the hind-brain, and when the paths connecting the subordinate centre in the lumbo-sacral cord with the hind-brain are divided, or the higher centres injured, the emptying of the bladder remains incomplete. The spinal centre can be excited reflexly by stimulation of any sensory nerve, such as the sciatic or the fifth nerve. In many cases where, in consequence of obstruction to the passage of impulses from the higher parts of the central nervous system, micturition is delayed, this act may be excited by the application of cold or hot sponges to the perineum, and it is well known that irritation of the pelvic organs in children may give rise to reflex involuntary micturition.

Barrington finds in cats that normal micturition is lost after destruction of a small area of the midbrain situated a short distance ventral to the internal edge of the superior cerebellar peduncle, and extending from the level of the posterior end of the aqueduct in front to the middle of the motor nucleus of the fifth nerve behind.

For Reference

GRUBER (1933). The Autonomic Innervation of the Genito-urinary System. *Physiol. Rev.*, **13,** 497.

LANGWORTHY, KOLB and LEWIS (1940). Physiology of Micturition. (London: Baillière, Tindall and Cox.)

CHAPTER XLVIII

THE SKIN AND THE SKIN GLANDS

In all classes of animals, the skin performs two functions. In the first place it serves to protect the underlying parts from injury and from invasion by foreign organisms. In the second place, it serves as a sense organ, and is richly supplied with nerves, by means of which the activities of the body as a whole may be brought into relation with the changes going on in the environment. In warm-blooded animals, the skin plays an important part in the regulation of the body temperature, since the loss of heat from the body must occur almost entirely through its surface. It also has important functions as a place for the storage of certain materials.

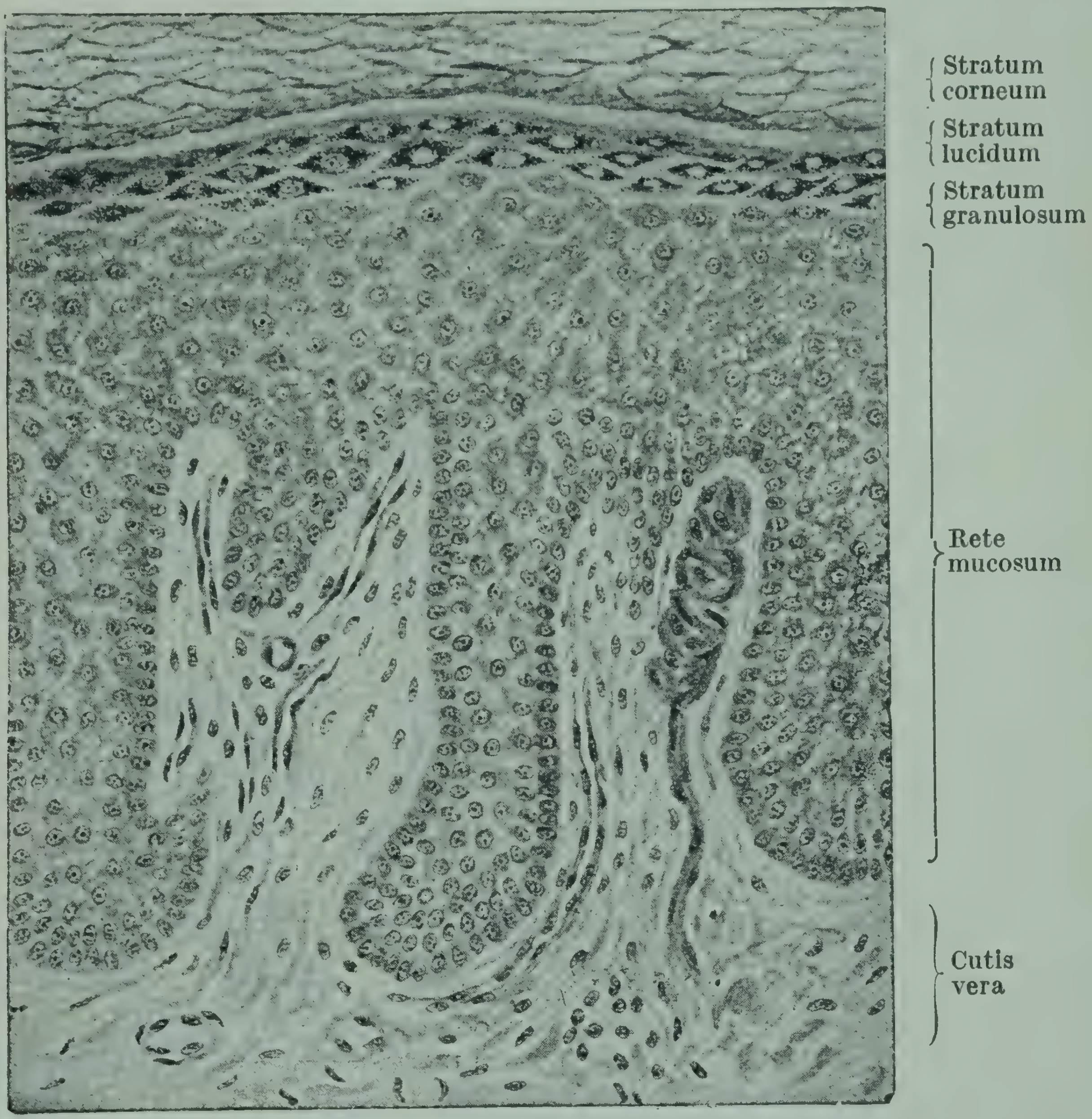


Fig. 629. Vertical Section through the Skin of the Palmar side of the Finger, showing two Papillæ (one of which contains a Tactile Corpuscle) and the deeper Layers of the Epidermis. Magnified about 200 diameters. (Schäfer.)

The development of the skin as an organ of protection shows wide modification in various classes of animals. Thus it may become the seat of formation of horny plates as in the alligator; of poisonous glands, as in the toad; or of mucous glands, as in many varieties of fishes. In warm-blooded animals, the development of hair from the deeper layers of the epidermis serves to diminish the loss of heat. Since the hair follicles are richly supplied with nerve fibres, the hairs act also as organs of sensation. In man, where the hairs are rudimentary, except in certain localities, practically only this tactile function is retained. The external layer of the skin in man consists of a tough horny layer formed by the keratinisation of the external layers of cells of the epidermis. The skin is composed of two parts, the epidermis and the cutis (Fig. 629). The epidermis is a stratified squamous epithelium. The deeper layers form the *rete mucosum*, being soft and protoplasmic, while the superficial layers forming the cuticle are hard and horny.

The deepest layers of the *rete mucosum* often contain granules of a pigment, *melanin* ; in the coloured races this is abundant, and it is also produced in white skins as a result of sunburn. The formation of melanin is due to the presence of an oxidase in the *melanoblast* cells ; fresh or formalised sections of skin rapidly form melanin in these cells if immersed in a solution of 3 : 4 dihydroxy-phenylalanine.[*] This substance is the first stage in the oxidation of tyrosine by the oxidase called tyrosinase,[†] and we may suppose that tyrosine is the parent substance of melanin.

The most superficial layer of the rete mucosum is formed of flattened cells filled with granules of a material staining deeply with hæmotoxylin and eosin, known as *eleidin*. This layer is called the *stratum granulosum*. Immediately superficial to this layer is another in which the cells are indistinct. The cells are clear in section and form what is known as the *stratum lucidum*. These two layers evidently constitute the intermediate stages in the transformation of the cells of the rete mucosum into the horny scales which make up the *stratum corneum*.

The cutis, or corium, is composed of dense connective tissue, which becomes more open in texture in its deeper part, where it merges into the subcutaneous connective tissue. The elasticity of the skin is due to the presence in the corium and subcutaneous tissue of numerous fine elastic tissue fibres. In old age this is diminished in amount, and the collagen fibres increased, especially where wrinkles occur. The most superficial layer of the corium is prolonged into minute papillæ, over which the epidermis is moulded. These papillæ contain, for the most part, capillary blood vessels ; a few contain touch corpuscles, special organs of tactile sensation. The blood vessels of the skin form a close capillary network immediately at the surface of the cutis, sending up loops into the papillæ. Direct anastomoses between arterioles and venules (arterio-venous anastomoses) are common in the skin in parts liable to become cold, *e.g.* the hands, feet and digits in man, the ear in rabbits, etc.[‡]

All parts of the skin, except the palms of the hands and the soles of the feet, are beset with hair follicles. The hair follicles are small pits which extend downwards into the deeper part of the corium, being downgrowths of the rete mucosum. The hair grows from a small papilla of cells at the bottom of the follicle, the part of the hair lying within the follicle being known as the hair root. The hair itself consists of long, tapering, horny cells, the nuclei of which are still visible, though the cell substance has been almost entirely converted into keratin.

THE SEBUM. The cuticle is kept supple and preserved from the drying effects of the atmosphere by being constantly impregnated with a fatty material known as *sebum*. This material is formed by the *sebaceous glands*, which are present wherever hair follicles are to be found, and to it the secretion of any local *apocrine* sweat glands is added. A sebaceous gland is a pear-shaped body, consisting of a secreting part and a short neck opening into the follicle. The gland proper is composed of a solid mass of cells. The outermost cells are flattened and generally show signs of proliferation. The cells lying internal to these are much larger, and their protoplasm is transformed into a network, in the meshes of which are granules which show the reactions of fat. Further inwards, the protoplasmic network diminishes in amount, while the fatty granules increase in size, so that, in the lumen adjoining the duct, we find only a mass of cell débris and droplets of fatty material. It has often been thought that the secretion of sebum depended simply on a fatty degeneration of the cells. The granules, however, when they first appear, stain with acid fuchsin, but not with osmic acid, and one must regard the formation of sebum as an act of true secretion, in which secretory granules are gradually transformed into the special constituents of the sebum. The fatty acids of the sebum are for the most part in combination, not with glycerol but with higher alcohols, including isocholesterol. Wool-fat or lanoline, which may be extracted from wool, as well as from the feather-glands of water birds, such as the goose and duck, consists mainly of esters of isocholesterol. It presents many advantages over ordinary fat as a protective salve for the surface of the body. It can take up as much as 100 per cent. of water ; it does not tend to become rancid ; it has a destructive action on streptococci, and other organisms, which it owes to certain fatty acids and soaps.[§]

The secretion of sebum is a continuous process, though it is probably quickened in conditions of increased vascularity of the skin. The extrusion of the products of secretion is determined by the presence of unstriated muscle fibres, the *arrectores pilorum*, which pass from the surface of the cutis obliquely over the outer surface of the sebaceous gland. They are supplied by sympathetic nerves. When these muscle fibres contract, the hair is erected and a certain amount of the sebum squeezed out on to the root of the hair and the surrounding skin. This contraction will occur on faradic stimulation, or whenever cold is suddenly applied to the skin. In some animals, including man, contraction also occurs in response to adrenaline.[||] The contracted condition of all the muscles of the hair follicles is shown by the " goose-skin " produced under such circumstances. There is no evidence that the secretion of sebum is in any way under the control of the central nervous system

---

[*] BLOCK. *Z. f. physiol. Chem.*, 1917, **98**, 226.

[†] RAPER. *Physiol. Rev.*, 1928, **8**, 245.

[‡] GRANT. *Heart*, 1931, **15**, 281, 385.

[§] BURTENSHAW. *J. Hyg.*, 1942, **42**, 184.

[||] ELLIOTT. *J. Physiol.*, 1904, **32**, 416.

The Sweat Glands. The sweat glands are distributed over the whole surface of the skin, and are especially abundant on the palm of the hand and on the sole of the foot. They are composed of single, unbranched, coiled tubes, which lie in the subcutaneous tissue and send their ducts up through the cutis, to open on the surface by corkscrew-like channels which pierce the epidermis. The secreting part of the tube consists of a basement membrane lined by a double layer of cells; the innermost of these are cubical and represent the secreting cells proper. Between the secreting cells and the basement membrane is a layer of unstriated muscle fibres. The duct of the gland has an epithelium, consisting of two or three layers of cells with a well-marked internal cuticular lining, but there is no muscular layer.

In addition to these, which are sometimes called the *eccrine* sweat glands, there are in certain situations, *e.g.* axilla, larger glands somewhat resembling sebaceous glands, and often, like them, associated with hairs. They are called *apocrine* glands and secrete a fluid containing a fatty material of peculiar odour. The *ceruminous glands* of the external auditory meatus are similar, and secrete a waxy material.

Insensible Perspiration. It has been known since the work of Sanctorius in 1614 that a person slowly loses weight between meals. Since loss of weight = loss of water + loss of CO_2 — gain of O_2, the net loss at usual values of R.Q. is chiefly due to loss of water by skin and lungs, and only to the extent of 10 per cent. or less, to the difference between oxygen intake and CO_2 output. It varies according to the metabolic rate, as would be expected, since it represents the means by which, in a majority of cases, about 25 per cent. of the body heat is lost.[*] It also varies with the temperature and humidity of the air. Benedict and his colleagues [†] estimate the loss at about 18 grams per hour for lungs and the same for the skin, in a man with a basal metabolism of 1,700 cals.: this would amount to about 430 grams from the skin and a like amount from the lungs in the twenty-four hours. In the more humid atmosphere of Europe it may be less. About one-quarter of the total loss from the skin takes place from the hands and feet, which lose moisture at about three times the rate found for equal skin areas elsewhere. These estimations refer, of course, to the condition in which there is no visible sweating, but only a loss of so-called "insensible perspiration" which is continually going on. It is apparently not due to slight sweating, but to a passage of water vapour by osmosis, through the skin,[‡] and, as it seems, not by way of the sweat glands at all.[§] Individuals lacking sweat glands lose as much by insensible perspiration, in cool surroundings, as normal controls do.

The Sweat, which mainly consists of the secretion of the eccrine glands, is the most dilute of all animal fluids. As collected, it generally contains epithelial scales and some sebum, and is usually acid in reaction; the *p*H, however, may range from 4 to 8, according to the extent to which the lactic acid is neutralised. The composition of sweat varies considerably, but the following results are typical:—

COMPOSITION OF HUMAN SWEAT.[§]

Water		99·742–99·221	per cent.
Solids		0·258–0·779	,,
Organic solids		0·03–0·290	,,
Ash		0·144–0566	,,
Chlorine (usually over 0·15 per cent.)		0·059–0·346	,,
Potassium	about	0·017	per cent.
Sulphate	,,	0·004	,,
Lactic acid	,,	0·070	,,
Urea		0·03	,,
Sodium	,,	0·15	,,
Sugar	,,	0·004	,,

---

[*] SODERSTROM and DUBOIS. *Arch. Int. Med.*, 1917, **19,** 931.

[†] BENEDICT and ROOT. *Arch. Int. Med.*, 1926, **38,** 1.

[‡] HANCOCK, WHITEHOUSE and HALDANE. *Proc. Roy. Soc.*, 1930, **105B,** 43.

[§] VASTI. *Amer. J. Physiol.*, 1932, **102,** 60.

[||] Based on WHITEHOUSE. *Proc. Roy. Soc.*, 1935, **117B,** 139.

The only plentiful constituent is sodium chloride, the amount of which varies, though sweat is always hypotonic. In profuse sweating a man may lose 10 g. or more of Cl in three hours. In winter, or in subjects unaccustomed to sweating, the chloride content is much higher (*e.g.* 50 mE. per l.) than in those acclimatised to heat or used to sweating (*e.g.* 25 mE. per l.). The more rapid the sweating the higher the chloride concentration. Urea is present in about the same amount as in blood, but lactate seems to be either concentrated or formed by the sweat glands. After filtration, sweat forms a clear colourless fluid of a specific gravity of about 1,003.

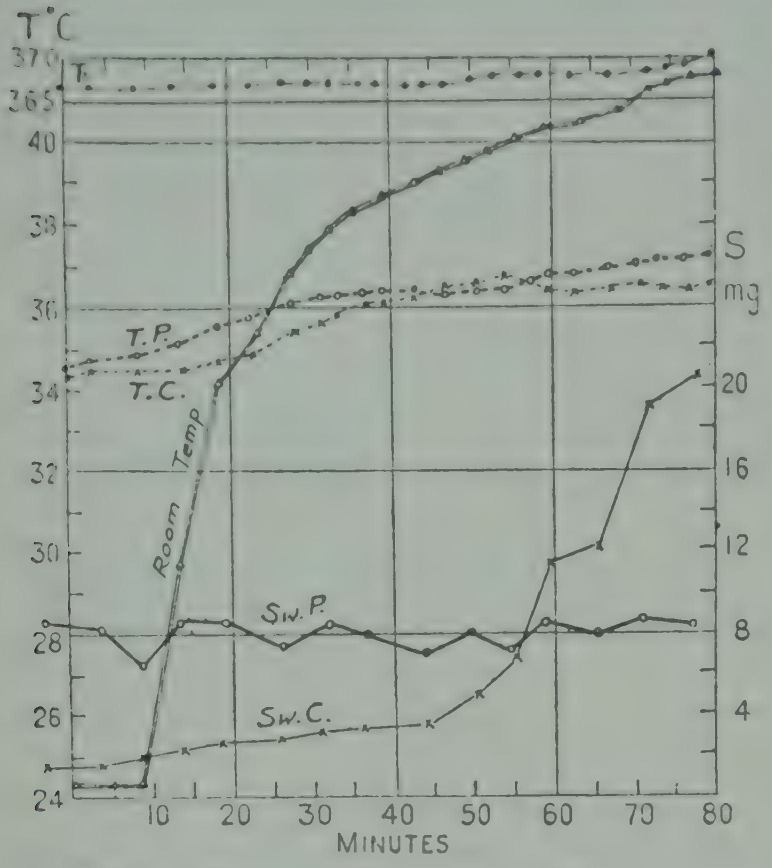


FIG. 630. Sweating in m.g. per 20 sq. cm. per 5 minutes from palm (Sw.P.) and chest (Sw.C.) with rise of room temperature. Tr. T.P., T.C. = temperatures in rectum, on palm and on chest respectively. (KUNO, *Lancet*, 1930.)

Sweating. Many sweat glands are inactive and incapable of secreting, but there are from 2 to 3½ million functional sweat glands in the body.* The excretion of sweat by the skin is mainly determined by the need for regulating the temperature of the body, so that the water lost in this way depends on the heat production of the body or on the external temperature, and is very little affected by alteration in the quantity of fluid drunk. If a man's body be enclosed in a vessel through which a current of air is passed, and the temperature of the air be gradually raised, it will be noted that the amount of water given off rises slowly up to a certain degree and then mounts rapidly. The sudden kink in the curve is due to the commencement of the activity of the sweat glands (Fig. 630), *i.e.* to onset of sweating.

The secretion is under the control of the central nervous system and, except for that of the palms and soles, has been shown by Kuno † to be almost entirely adapted to the regulation of the body temperature. In a hot humid climate a lightly clad man can sweat 200 g. per hour when at rest, and at work up to 900 g./hr. The nervous mechanism can be set in to activity either centrally or reflexly. The most usual cause is a rise of the body temperature. If a man sits in a warm room, *e.g.* in a Turkish bath, the secretion of sweat commences at a mean skin temperature of 34–35° C., or as soon as the mean temperature of the body has attained a height of 0·2° to 0·5° C. above normal. Sweating is not uniform in different parts of the body. The palms and soles do not react to the rise of temperature ; in profuse sweating about 50 per cent. of the total sweat is given by the trunk, 25 per cent. by head and upper limbs and 25 per cent. by the lower limbs.‡ It will be seen that most sweating occurs in those regions which, by reason of their exposure, offer the best opportunities for evaporation.

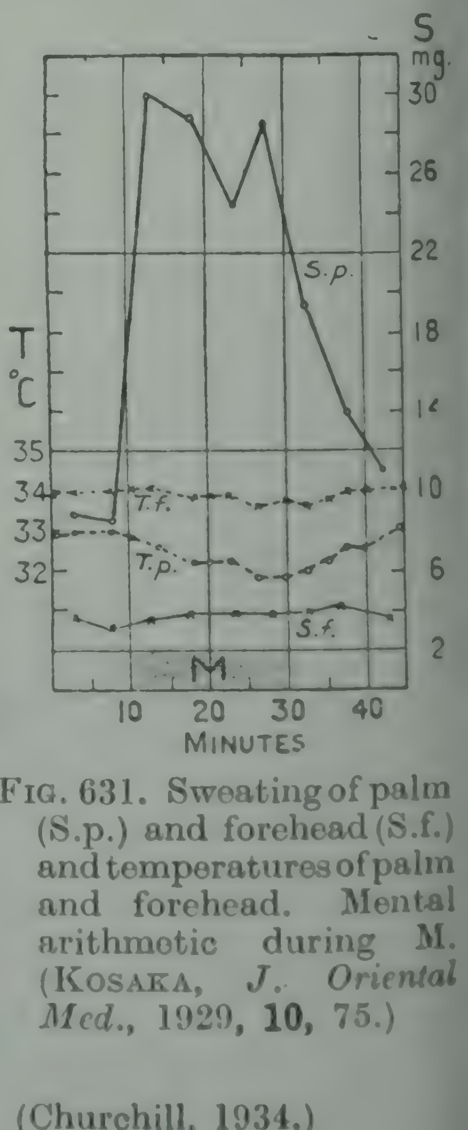


FIG. 631. Sweating of palm (S.p.) and forehead (S.f.) and temperatures of palm and forehead. Mental arithmetic during M. (KOSAKA, *J. Oriental Med.*, 1929, **10,** 75.)

---

* KUNO. *Lancet*, 1938, **234,** 299.

† KUNO. "The Physiology of Human Perspiration." (Churchill, 1934.)

‡ WEINER. *J. Physiol.*, 1945, **104,** 32.

The palms and soles are peculiar inasmuch as they constantly secrete some sweat, but the amount is not increased by rise of temperature (Fig. 630). It does, however, show an increase under direct nervous control, as in exercise, anxiety, or mental stress of any kind (Fig. 631). This probably applies also to the pads of the cat, which cannot play an important part in the regulation of body temperature, since they are the only parts of its body provided with sweat glands.

Sweating is facilitated by muscular exertion, and the body temperature will generally be found to be raised if it be taken at the instant that sweating has commenced. The effect of rise of temperature may, however, be partly local, so that one arm enclosed in a hot-air bath may sweat more than the rest of the body. Under ordinary circumstances, the central stimulation by the warm blood is the predominant factor. Sweating may be induced in the cat's paws by warming the blood passing through the carotid arteries on its way to the brain, at a time when the temperature of the blood circulating through the rest of the body, including the paws themselves, has undergone no alteration. Sweating may also be aroused by asphyxia, and this result is found even in the spinal cat, *i.e.* after separation of the spinal centres from the medulla. The secretion of sweat, although generally associated with increased vascularity of the skin, is not entirely dependent thereon. Thus, even in the amputated limb, stimulation of the sciatic nerve may cause the appearance of drops of sweat on the pad of the foot. If the sciatic nerve be stimulated in the intact animal, the secretion of sweat which is produced is associated with constriction of the vessels of the skin, due to simultaneous stimulation of the vaso-constrictor fibres running in the sciatic nerve.

As Langley showed,* the nerves to the sweat glands are sympathetic. Leaving the cord by the white rami from the second dorsal to the third or fourth lumbar nerves, they pass into the sympathetic chain. Here the first relay of fibres ends in connection with the cells of the sympathetic ganglia, and a fresh relay of fibres, which are non-medullated, pass from the cells along the grey rami into the various spinal nerves, to be distributed to the whole surface of the skin. There are certain anomalies connected with sweat nerves, however. Thus adrenaline has, in man, no action, whereas pilocarpine, ¼ gr. (16 mg.) or mecholyl, 25 mg., which are parasympathetic excitants, cause abundant sweating, even after division of the sweat nerves, so that they must act peripherally on the glands. Complete section and degeneration of all the nerves to a limb, however, greatly reduces and finally abolishes the response to pilocarpine, and it also reduces the vasodilator response (Burn).† The action of pilocarpine, as well as the effects of artificial stimulation of the sweat nerves, is abolished by the administration of atropine. We may therefore suppose that, as in the case of the sympathetic ganglion synapse, the "chemical transmitter" in the case of the sweat glands, is not adrenaline, but acetyl choline, which substance does, in fact, cause sweating. The injection of pilocarpine or pituitrin into the cerebral ventricles also causes profuse sweating,‡ apparently by exciting the diencephalic centres connected with the origins of both sympathetic and parasympathetic systems.

Absorption by the Skin. In order to test the alleged influence of baths containing medicinal substances in solution, many experiments have been made to determine whether absorption is possible by the skin. It may be regarded as established that the uninjured skin is impermeable to watery solutions of salts or other substances, though ions can be carried through by an electrical current (ionophoresis). On the other hand, it is possible to produce a certain amount of absorption by the application of substances conveyed in fatty vehicles. Thus, the administration of mercury is often carried out by the inunction of mercurial ointments, and the fact that mercurial salivation may be produced in these conditions shows that a certain amount of the mercury must have been absorbed.

---

* LANGLEY. *J. Physiol.*, 1891, **12**, 347.

† BURN. *J. Physiol.*, 1922, **56**, 232; 1925, **60**, 365.

‡ CUSHING. Papers relating to the pituitary body, hypothalamus and parasympathetic nervous system. (Baillière, Tindall and Cox, 1932.)

It is difficult to imagine that any appreciable amount of cod liver oil will be available for the nutrition of the infant when this substance is administered by rubbing it on the skin. On the other hand, the moist mucous surfaces, such as the conjunctiva or the mucous membrane of the respiratory passages, as well as raw surfaces of the skin, *e.g.* which have been deprived of their epidermal layer by the application of blisters, permit the rapid passage of substances in either watery or oily solution.

Gaseous Exchanges through the Skin. In any animal with a thin moist skin, such as the frog, the absorption of oxygen and the excretion of CO_2 through the skin may be sufficient for the proper aeration of its blood, so that it may continue to live after the extirpation of its lungs. In man there is also a continuous output of CO_2 through the skin, but the amount leaving the body in this way is negligible compared with that which is exhaled through the lungs. The loss of CO_2 by the skin rises with increase of external temperature. Thus at a temperature of 29° to 33° C. the CO_2 output by the skin is about 0·35 g. per hour, *i.e.* about 8·4 g. in the twenty-four hours. When the external temperature rises above 33° C., the CO_2 output increases, so that at 34° it is doubled and at 38·5° it may amount to as much as 1·2 g. per hour (Schierbeck). It is just at this temperature of 33° C. that a secretion of sweat begins to be noticeable, so that it seems probable that the increased CO_2 output may be due to the secretion of sweat, which, like other body fluids, contains carbon dioxide that readily escapes from it on exposure to the air.

Storage Functions of the Skin. The corium, together with the subcutaneous tissue of the skin, has considerable capacity for storing various materials, provisionally or permanently. Of these the best known is fat, which is laid down in fat cells as a permanent store of subcutaneous adipose tissue.

The other type of storage which takes place in the skin is what has been called by Cannon *storage by inundation*, and which may be compared with the storage of water in a bog. The areolar tissue of the skin, and other organs, readily take up more water, and is an important place for its storage. Excess water retained in the body is held most conspicuously in the subcutaneous tissues, and, in pathological states, may lead to œdema ; conversely, in water deprivation the subcutaneous water stores are freely drawn upon. After severe hæmorrhage, this may happen very quickly, about half the water which leaves the tissues to enter the blood coming from the skin, the remainder chiefly from the muscles.

When the blood is flooded with soluble substances, these may likewise find temporary storage, accompanied by water or not, in the subcutaneous tissues. The most noteworthy instances are of glucose and salts. When the blood sugar level rises suddenly, considerable amounts of sugar may find temporary lodgment in the subcutaneous tissues, and as the heavy load of sugar is removed from the blood for utilization, or storage as glycogen, so the flood of sugar subsides from the connective tissues into the blood again. It is similar with chlorides. When a chloride-rich diet is given, or after intravenous chloride infusion, the heaviest storage of chloride is in the skin, while in chloride starvation, in which the blood chloride is preserved almost unchanged, as much as 90 per cent. of the chloride lost from the body may have come from the skin, though this tissue forms only 16 per cent. of the body weight.

THE HAIR

Hair in the lower animals has definite functions in relation to heat regulation, but in man this is of slight importance. The beard is a male secondary sexual character, as is hair on the chest and abdomen, and is not developed in men castrated before puberty. The hair of the axillæ and pubic regions is a secondary character of both sexes and is undeveloped after early removal of the gonads. There are considerable racial variations in hair.

There is no doubt that the growth and distribution of hair follicles is affected by hormones and vitamins, but exact knowledge of the effects is not yet available. The suprarenals seem to exert an influence, as in suprarenal tumours hair growth may be abnormal, but since sexual precocity is also a feature, an influence through the gonads may be more fundamental. The thyroid hormone is also essential to the proper growth and normality of the hair.

Curliness or straightness has not been satisfactorily explained ; the former belief that hairs curled because they were elliptical in cross section instead of circular appears to be inadequate to explain it. Artificial curling, as in "permanent waving," is due to elongation of some of the filamentous keratin molecules by stretching under heat (*vide* p. 43).

Pigmentation, due to deposition of melanin, may be diffuse, as in red hair, or granular, as in black. The pigment tends to bleach in sunlight.

Hairs do not grow indefinitely, but after a period of growth the length remains stationary and the hair sooner or later may fall, or be pulled, out, after which a fresh cycle begins. Each hair follicle has its own characteristic growth rate and cycle ; on the average the head hair grows about 0·4 mm. daily and normally reaches 65 cm. in length, if uncut, *i.e.* growth continues for about four years. Short hairs such as those on the forehead, grow more slowly. Shaving off the hair has not been proved to modify the rate of growth. Increased rate of growth is preceded by an increased metabolism in the hair bulb.*

---

* BUTCHER. *Amer. J. Physiol.*, 1943, **138**, 408.

For Reference

CALVERY, DRAIZE and LAUG (1946). The Metabolism and Permeability of Normal Skin. *Physiol. Rev.*, **26**, 495.

DANFORTH (1939). Physiology of Human Hair. *Physiol. Rev.*, **19**, 94.

KUNO (1934). The Physiology of Human Perspiration. (London: Churchill.)

NEWBURGH and JOHNSTON (1942). The Insensible Loss of Water. *Physiol. Rev.*, **22**, 1.

CHAPTER XLIX

THE TEMPERATURE AND HEAT-BALANCE OF THE BODY

The temperature of the body of cold-blooded, or *poikilothermic*, species differs little from that of the surroundings. Their torpidity in cold surroundings, and the risk of heat damage in hot ones are hence considerable handicaps to them.

The birds and mammals, by contrast, are called *homoiothermic* animals: they have body temperatures which vary but slightly, even in the face of widely different surrounding temperatures. The actual mean body temperature differs according to species, *e.g.* in birds it is usually 39° to 40° C., and in mammals 37° to 40° C. Thus in man it is 37°; in the horse 37·8°; in the cow, cat and dog 38·6°; in the rabbit 39°; the pig 39·2°, and in the sheep 40° C. Their temperature is higher than that of their usual surroundings, and this is owing to the heat generated by oxidative processes continually going on within the body: in hot surroundings, however, the body may, for a time, be cooler than its environment, so there must be some active method of cooling.

The approximately even temperature is obviously due to the existence of a more or less perfect state of balance between heat production and heat loss. Since heat production = heat loss, which is approximately equal to heat leaving the skin, it is evident that all these values can commonly be related to the surface area of the body.* Accordingly they are expressed in terms of kilocalories per square metre of body surface per hour (Cal./sq. m./hr.)† as in the case of B.M.R. measurements.

It takes a smaller heat-exchange to warm or cool the body than to warm or cool an equal weight of water, because the specific heat of the body is only about 0·83; hence a man of 70 kg. has only the same thermal capacity as about 58 kg. of water.

The Body Temperature in Man. It should be understood that the idea of a stated body temperature (T_b) is somewhat of an abstraction, since the temperature varies from point to point in or upon the body. Active tissues, such as liver, or contracting muscle, have higher temperatures than inactive tissues, or than tissues which are losing heat rapidly, as the skin does. But, owing to the circulation of the blood, and to the conductivity and thermal capacity of the tissues, there is a constant tendency for the temperatures of the interior of the body to equalise, and for there to be a gradient of varying steepness from the deep tissues to the skin. This gradient is illustrated by the following temperatures.

SURFACE

Skin covered with clothes	33° to 35° C.
Naked skin in bath at 5° C.	17° C.
,, ,, ,, 25° C.	26·5° C.

MUSCLES 12 MM. BELOW THE SURFACE

In bath at 5° C.	36·3° C.
,, ,, 25° C.	36·9° C.

It is common practice to take the temperature of man in the rectum, or mouth, or occasionally in the axilla. The rectal temperature lies nearest to the mean deep body temperature: that in the mouth varies with rate of breathing, and with loss of heat through the cheeks: in determining it the thermometer should be left *in situ* beneath the tongue for (usually) at least three minutes, or until the maximum is reached. A thermometer placed in the axilla takes several minutes to reach its maximum reading. Usual temperatures are: rectal, 37·5° C.; mouth 37·0° C., and axilla 36·5° C., in healthy adults, resting, and well clad. It will be seen that the prevalent unqualified statement that the "normal temperature" is 98·4° F. (36·9° C.) is meaningless: many young adults, when up and at rest have temperatures below this, usually about 97·5° F., on the average during the day; fluctuations are definite, usually of at least ± 1° F., during the day. The highest temperature is usually at about 6 p.m., and the lowest at about 4 a.m., and the oxygen usages show parallel variations. They are probably caused by variations in the state of tension in the muscles, and in the activity of digestion. If the habits of a man be reversed, so that he sleeps by day and works at night, it is possible partially to reverse the diurnal rhythm of temperature.[*] The temperature may also be transitorily affected by various acts, such as the taking of hot or cold baths,[†] eating ice, taking hot food, or muscular work,[‡] the last-named often producing considerable rise. The temperature is often higher in females (except during eight to twenty-one days after menstrual periods) and in young individuals than in males or older ones.

---

* BENEDICT. "Vital Energetics." (Washington: Carnegie Inst.)

† 1 Cal. = 1,000 cal. = roughly 4 B.Th.U. or 0·00004 Therm.

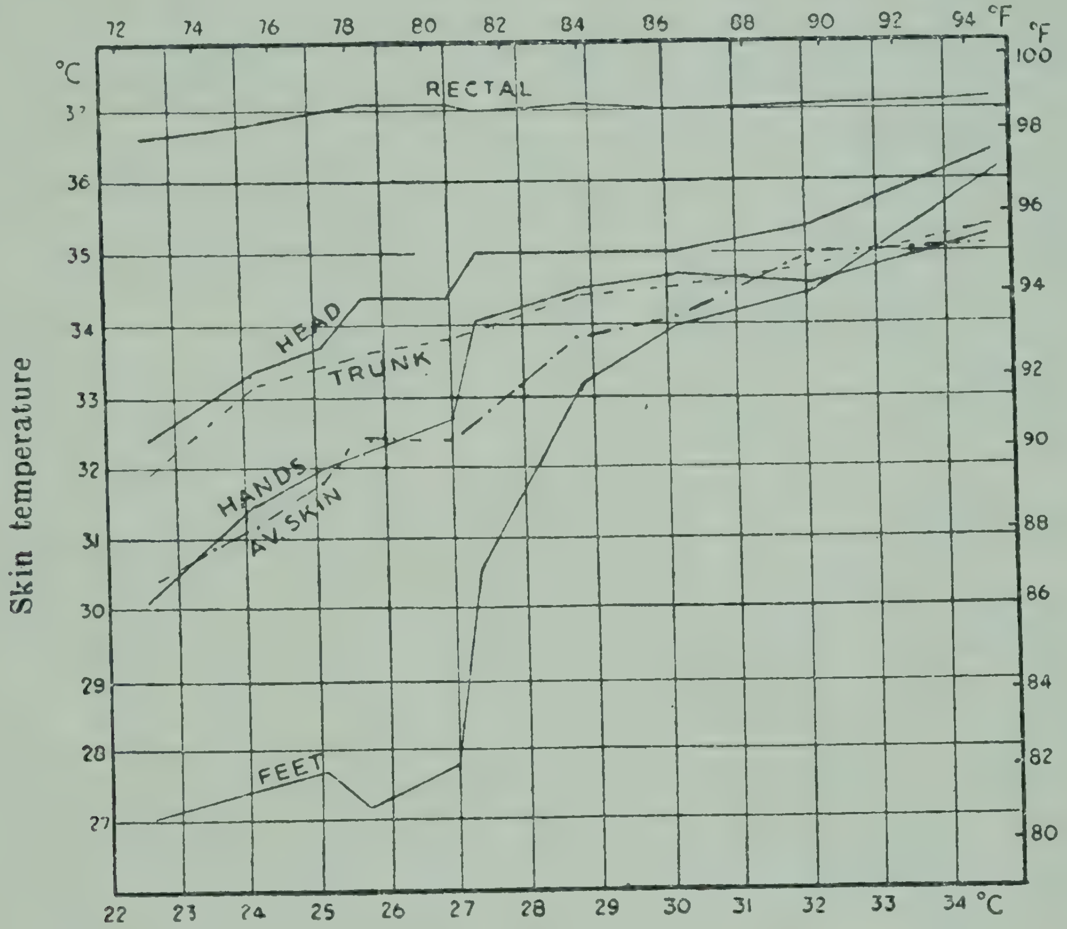


FIG. 632. Rectal and Skin Temperatures at Different Places, and mean temperature, after 1½ hours at varying Air Temperatures. (BRUNT, after HARDY and DU BOIS, 1941.)

There are limits to the power of the organism to regulate the body temperature even approximately when faced with excessive heat production from within, as in strenuous muscular exercise; or with greatly increased rate of heat loss, as in cold baths, arctic climates, or deprivation of clothing; or with increased production and much reduced rates of heat loss, as in exercise in hot moist surroundings. In these circumstances the body temperature may rise or fall considerably. Regulation is more imperfect in infants than in adults.

The Temperature of the Skin. Skin temperatures (T_s) can be measured by means of thermo-electric junctions, or, more accurately, by the Hardy radiometer. The temperature varies considerably according to position, state of local circulation, sweating, climatic conditions, clothing, etc. Beneath adequate clothing the trunk skin temperature is usually about 33° C.: with nude subjects the head and neck are warmest (33° in room at 24°), and feet the coldest (27°) (Fig. 632).

---

[*] BENEDICT. *Amer. J. Physiol.*, 1904, **11**, 145.

[†] DILL and FORBES. *Amer. J. Physiol.*, 1941, **132**, 685.

[‡] MACKEITH, *et al.* *Proc. Roy. Soc.*, 1923, **95B**, 413.

The *mean skin temperature* is the average temperature taken at 15 or more points suitably distributed on the skin surface. It rises about 1° for each 2° rise in surrounding temperature, until the latter reaches 28°–30°, after which it remains level for a while. The "comfort zone" of *mean* skin temperature is 30°–35° C. The *mean temperature of the body* can be calculated from the rectal and mean skin temperatures. These data are valuable in dealing with the physical aspects of heat balance.

REGULATION OF BODY TEMPERATURE

The small variation in the body temperature of mammals connotes an equalization of heat production and heat loss by adjustment of one or both. It will be convenient to deal under separate headings with these two methods of regulation. But first a prevalent misconception must be cleared up. This is the belief that a raised temperature necessarily means increased heat production or diminished heat loss, or both.

A *rising* temperature means that the rate of heat production is greater than that of heat loss, and conversely for a falling temperature. But a steady *raised* temperature means that once more heat production and loss are balanced, though the regulator is, for some reason, set to a different level; the raised temperature by itself tells us nothing about the rate of production and loss, except that they are equal when it is constant.

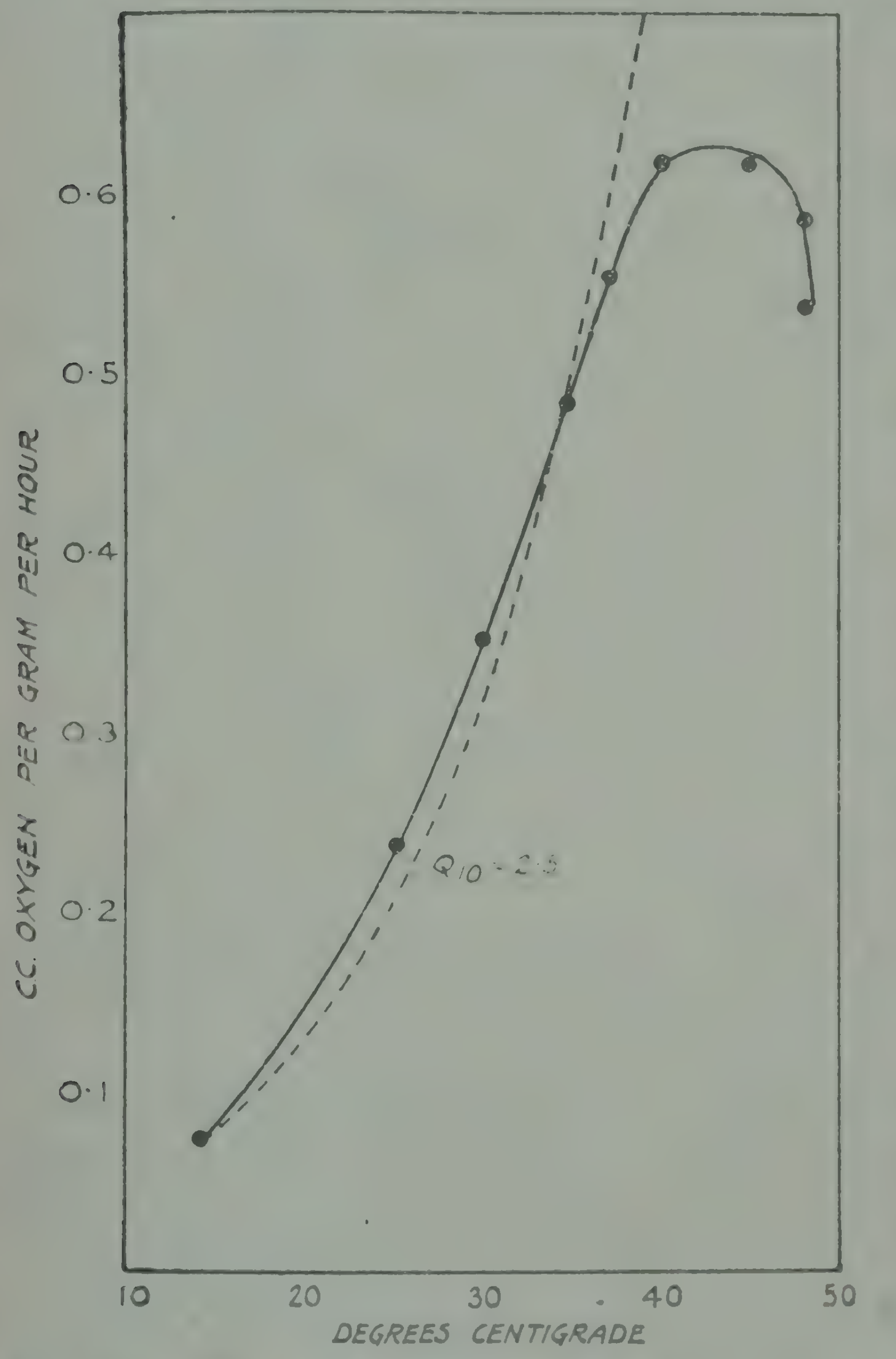


FIG. 633. Effect of Alteration of Temperature on Oxygen Usage of Plain Muscle (Guinea-pig Uterus). Abscissæ = temperature. Ordinates = c.c. oxygen per gram per hour. The dotted line represents the theoretical curve if the oxygen intake were increased two and half times for a rise of 10° C.

The Regulation of Heat Production. In the oxidative processes responsible for heat production all the tissues participate. In the regulation of heat production the muscles probably play the leading part: even during rest these are the seat of a restricted heat production, which is largely increased with every active contraction and with heightened states of tonus. Weight for weight, the liver is probably one of the most constant and active sources of heat: the temperature of the blood of the hepatic veins is often the highest in the body. If the oxygen utilisation of the body is increased, the heat production is also raised, and if the heat loss is not also proportionately increased the body temperature will be increased until loss again equals production. Thus the intravenous injection of 0·02 g./kg. of dinitro-α-naphthol may cause the body temperature of dogs to rise to 45° C. and of pigeons to 50° C.; the oxygen consumption in such cases may be increased several-fold.*

When isolated tissues are subjected to alterations of temperature, their oxygen utilisation, and therefore their heat production, increases with rise, and decreases with fall of temperature, a rise of 10° C. increasing the oxidative processes two- to three-fold, as for many chemical reactions (Van't Hoff's law) (Fig. 633). Above a certain temperature (40°–45° C.) damage is done to the tissues, and the oxidative processes fall off again. The behaviour of intact cold-blooded animals is much the same: the change in their metabolism with change of temperature therefore tends to augment, rather than to counteract, the effects of environmental temperature upon the body temperature.

Since the homoiothermic animals maintain an (almost) even temperature despite very wide variations in their environment, we have to enquire whether, in cold surroundings at all events, there may not be some mechanism (quite apart from that brought into play by increase of voluntary movement) by which the rate of metabolism is adjusted, so that increased tendency to cool is met by an increased rate of heat production; the opposite response, of reduced metabolism when surrounding temperatures are raised, might also enter into the account. These responses do, in fact, occur, and constitute what has been called the *chemical regulation* of temperature: it is purely a regulation of production, and not of loss, of heat. But if, for any reason, the various regulative processes *fail* to prevent a change of body temperature, then the tissues will behave like isolated tissues, and so produce more heat the warmer they become, and *vice versâ*.

The heat production in the muscles is dependent on their connection with the central nervous system. If this connection be severed, either by curare or by section of the cord, or if the reflex play of impulses on the muscles be abolished by anæsthetics, the animal will react like a cold-blooded animal.† The total metabolism of the body and the total production of heat then sink to a minimum, and are diminished by application of cold, or increased by application of warmth. On the other hand, in the intact mammal, moderate changes of temperature in the environment reflexly provoke changes in the *opposite* direction. Thus, exposure to cold below a critical temperature of about 15° C. for the nude subject, increases the heat production of the body.

The effects of variations in the external temperature on the metabolism of warm-blooded animals are well shown in the experiments of the Tables given below on the oxygen usage of a rabbit and of a curarised dog.

Rabbit (Pflüger).		Curarised Dog (Krogh).	
Rectal temperature °C.	c.c. Oxygen per Kg. and hour.	Rectal temperature °C.	c.c. Oxygen per Kg. and hour.
22	457	14·1	134
26	608	22·7	300
37·5	888	28	495
37·6	839	32·3	620
38	763	37·2	812
38·8	738	39·9	700

In the rabbit, cooling even to 26° only reduces the metabolism by 18 per cent. of that at 38·8°, while at 37·5° (a lowered body temperature), the metabolism is actually increased by some 20 per cent. The dog, deprived of its heat-response by curare, however, shows a fall of metabolism at all temperatures below that normal for the species, and also at the higher one.

---

* Heymans and Bouckaert. *Arch. inst. Pharm.*, 1928, **35**, 63; Euler, *ibid.*, 1932, **43**, 67.

† Pflüger. *Pflüger's Arch.*, 1878, **18**, 305; Kennaway and Pembrey. *J. Physiol.*, 1912, **45**, 82.

The immersion of a man in a cold bath for a minute or two considerably increases his oxygen usage, but his temperature may show little change. It is possible, however, by an act of the will, to keep all the muscles in a state of relaxation. Under these circumstances the temperature of the body rapidly falls, and with it the rate of metabolism.

Nude men or women, lying down in surroundings below 29° C., lose heat faster than they produce it, and soon respond by a rise of metabolism. In an environment between 29° and 30° heat loss and production are equal in both sexes, and up to that point their B.M.R.'s are equal. Between 30° and 32°, however, the B.M.R. of women falls to 14–20 per cent. below that of men: women, therefore, show a better chemical regulation of body temperature, and hence sweat less, in hot surroundings than men.* Moreover, the thick layer of subcutaneous fat insulates them more from cold, so that their skin temperatures may be 1° C. lower than in men, by which they lose less heat; in higher temperatures, on the other hand, the flushing of the skin raises its temperature 1·7° above that of men, which gives them further advantage.

SHIVERING. A much more vigorous response to cold is that due to increase of general muscular tone, and ultimately to shivering. For example, a dog exposed to sudden cold may only show a 7 per cent. increase in oxygen usage prior to the onset of shivering, but as much as 30 per cent. increase when shivering sets in.†

Though temperature changes no doubt affect the state of activity, and the tone of the muscles, and cold may also produce shivering, moderate exposures to cold lead to some increase of metabolism by other means. Probably a discharge of adrenaline from the suprarenals is one of the chemical factors which brings about an increase in general metabolism under these circumstances.‡ It has been found that the basal metabolism is considerably raised by adrenaline injection §; after removal of both suprarenals the body temperature is lowered and regulation is less precise. Further, certain drugs, *e.g.* tetrahydro-β-naphthylamine, which cause great rise of body temperature, evoke general sympathetic excitation and discharge of adrenaline from the suprarenals ||; again, exposure to cold causes reduction of the adrenaline content of the suprarenals.¶ Nitro-bodies, however, cause a rise of temperature after removal of the suprarenals, and must act directly on the tissues. An increased activity of the thyroid has also been claimed to result from prolonged exposure to cold **; this would also raise the metabolic rate.

This adjustment of the body temperature by variations in the heat production is extravagant of energy if the difference in temperature between animal and environment is considerable. In man it is mainly a response to cooling. But in the very good adjustment of the temperature which is present in the higher mammals, regulation of the heat loss plays a greater part than regulation of heat production.

The Regulation of Heat Loss. In warm-blooded animals the regulation of heat production is dominated by physiological factors; the control of heat loss is the resultant of closely interwoven phenomena, some physical, some physiological. It is convenient, as a first approximation, to consider these separately.

---

* HARDY and DU BOIS. *Proc. Nat. Acad. Sci. U.S.*, 1940, **26**, 389.

† HEMINGWAY and HATHAWAY. *Amer. J. Physiol.*, 1941, **134**, 596.

‡ HARTMANN and HARTMANN. *Amer. J. Physiol.*, 1923, **65**, 612; CANNON *et al.* *Amer. J. Physiol.*, 1927, **79**, 466.

§ BOOTHBY and SANDIFORD. *Amer. J. Physiol.*, 1919, **51**, 200; AUB *et al.* *Amer. J. Physiol.*, 1922, **61**, 349.

|| MUTCH and PEMBREY. *J. Physiol.*, 1911, **43**, 109; ELLIOTT. *J. Physiol.*, 1912, **44**, 379; CRAMER. *Brit. J. Exp. Path.*, 1920, **1**, 31.

¶ CROWDEN. *J. Physiol.*, 1929, **68**, 313.

** RING. *Amer. J. Physiol.*, 1939, **125**, 244.

Physical Factors. The modern approach, by the appraisal of the purely physical aspects, owes much to the work of Du Bois and Hardy, and of Gagge, Herrington and Winslow.

Heat loss from the body can occur by Conduction, Convection, Radiation and Evaporation. The relative part played by each will vary with such conditions as clothing, posture, temperature of air and walls, movement and humidity of air, etc. For simplicity, let us first consider the nude body, at rest in a room with walls and air at equal, and moderate, temperatures.

Heat is produced, chiefly in the deeper tissues, and is diffused by *conduction* to the cooler parts, and ultimately to the skin, in accordance with the relation :

$$\left.\begin{array}{c}\text{Total heat flow to skin}\\ \text{(in Cal./sq.m./hr.).}\end{array}\right\} = \left\{\begin{array}{l}\text{Conductance of superficial tissues (in Cal./sq.m./hr./°C.)}\\ \qquad\qquad \times \text{ Temp. gradient to skin } (T_b - T_s).\end{array}\right.$$

The value for minimal conductance of the superficial tissues is about 11.

A descending temperature gradient, of varying steepness in accordance with the varying conductance of the tissues, will evidently obtain from within outwards. Since epithelium and fat have a low thermal conductance (say, about ·00047–·0005, which compares well with cork = 0·0007), the gradient through tissues of and beneath the skin is especially steep, the maximum gradient being about 8° C. in 2·2 cm. : in some places a gradient may be detected to 10 cm. depth. This simple picture is apt to be modified by the circulation of the blood, which tends to even out such differences of temperature, so that the gradient may be reduced to as little as 1° C. over 2·2 cm.

Conductivity in most fluids is low (air = 0·00005 ; water = 0·0013 ; copper = 1·0), and is obscured by convection. Hence conduction proper only occasionally accounts for *loss* of heat from the body, as, *e.g.* when one sits on a cold marble slab (conductivity = 0·007).

Convection is much more important. Owing to the low specific heat of air (0·24 cal./g.), the temperature of air in contact with the skin rapidly rises ; in free convection a thin layer of air in contact with the body is moistened and warmed, thereby becoming lighter ; it rises, and its place is taken by cooler, drier air. The flowing layer of air is only 1 to 2 mm. thick : a layer next the skin clings to it and remains warm, so that the gradient from skin to air over 0·5 mm. is only about 1° C.

As a result of these convection currents, heat is continually being conveyed from the body surface to the air, at a rate proportional to the difference between their temperatures. At air temperatures of 34·5° C. the convection loss becomes zero, and at temperatures over 37° C. it becomes negative, *i.e.* the body is warmed by convection. If the surrounding air is in movement (or the body in movement through the air) the convection effect increases, and except at high speeds varies roughly as the square root of the air speed (usually $C = 16v^{\frac{1}{2}}\ (T_s - T_a)$ where v = air speed in m./sec.) ; it shows no further increase at speeds above 30 m./sec. (say, 70 mi./hr.).

Radiation is also an important source of heat loss. Human skin, of whatever colour, behaves in regard to long wavelengths as a perfect " black body," *i.e.*, it radiates strongly, and at a rate proportional to the fourth power of its absolute temperature (Stephan's law) : its radiations have a wavelength of 5 to 20 μ, with a peak at about 9 μ. The body surface, therefore, exchanges heat with its surroundings, such as walls, at a rate proportional to the difference between the fourth powers of their temperatures. If the temperature of the walls falls considerably, therefore, the radiation loss for a given skin temperature rapidly increases : but when the difference of temperature between walls and skin is not great the radiation loss may be taken as being approximately proportional to the difference between their temperatures (usually $R = 3{\cdot}6\ (T_s - T_w)$). It is, of course, only the area of projection of the body that radiates (it is often about 70 to 80 per cent. of the total body surface), and, therefore, the radiation loss will vary with posture, and from subject to subject.

Combined Effects of Radiation and Convection. When the difference in temperature between a warm object and its cooler surroundings is not large, the object cools, by radiation and convection together, at a rate proportional to the difference between their temperatures (Newton's law). From the physical properties of the object, and of its surroundings, which are expressed jointly in the form of a cooling constant (k), and from the differences of temperature, the actual loss of heat by radiation and convection can be calculated according to the expression $H = k\,(T_1 - T_2)$.

So long as the air is still, and the temperatures of air and surroundings identical, the human body also loses heat by radiation and convection in the same predictable manner, *i.e.* $R + C = K_o(T_s - T_a)$, and the "*environmental constant*" K_o for the nude is roughly $3{\cdot}6 + 16v^{\frac{1}{2}}$, which in nearly still air, has a value of about 8·45, where heat loss is in Cal./sq.m./hr. But when the air is warmer or cooler than the walls, or is in considerable motion, the problem is not so simple, for we cannot know what to consider as the "surrounding temperature," whether that of the walls to which the body loses by radiation, or that of the air, to which it imparts heat by convection. A solution was worked out by Gagge, Herrington and Winslow, by the use of an abstraction in the form of a weighted mean, called the "*operative temperature.*" This gives us a measure of the physical demands of the environment on the body, so far as radiation and convection are concerned, and the sum of these is predictable if the skin temperature (T_s) is also known. The standard operative temperature of a given place, when walls and air are at the same temperature, and with standard air movement of 5 m./min., is its dry bulb temperature: operative temperatures (T_o) for other conditions are deduced so as to give an imaginary environment with air and walls at equal temperatures and air movement at 5 m./min., and in which losses of heat by $R + C$ would be the same as in the actual environment. If T_s and the heat loss are known, $T_o = T_s - \dfrac{\text{heat loss to environment}}{5{\cdot}2}$.

Evaporation (E) contributes, according to conditions, to a very variable extent to heat loss. At environmental temperatures below 25° C., resting individuals do not sweat, and the evaporative loss, which is proportional to the metabolic rate, is entirely due to loss of water vapour from skin and lungs as "insensible perspiration": little sweating occurs at temperatures below 28°, but at 29° air temperature it definitely begins, and becomes of increasing importance with rising temperatures, so that in air at and above 35° C. *all* the body heat is lost by evaporation. The mean T_s at which sweating begins is about 34°, and it rises sharply at 35°: body temperature is also a factor that controls sweating, which is generally profuse at T_b of 37·2° C. or over.

When sweating sets in, the insensible perspiration probably ceases. Whatever the origin of the water *evaporated*, the heat abstracted from the surface of evaporation is 0·58 Cal./g.; if sweat runs off without evaporation it has, of course, no cooling effect. It is self-evident that if the relative humidity of the air is low, there will be a greater difference in vapour pressure between the skin (P_s) and the air (P_a) than when P_a is high, and, in fact, the rate of evaporation of the sweat can be shown to be proportional to ($P_s - P_a$), and also to increase with the speed of movement of the air (μ). The total amount of water evaporated will obviously vary with the area of skin which is wetted with sweat. The total vapour loss can be determined by weighing the body, and it is convenient to calculate from this, and from P_s and P_a, a figure designated $W\mu$, which $= \dfrac{E}{P_s - P_a}$, and is a measure of the area of wetted skin from which evaporation of moisture, sensible or insensible, may be supposed to be taking place at a given air speed. An alternative method of expressing the evaporative conditions is to calculate the *relative humidity of the skin*, (rh_s),* which depends on diffusion of water through the skin, the skin temperature, the relative humidity of the atmosphere, and air movement.

---

* Mole. *J. Physiol.*, 1948, **107**, 399.

The Heat Balance and its Partition. We may now consider the balance of heat production and heat loss of the body, and the extent to which the various factors contribute to the loss. The allocation of the different fractions of heat loss to their respective factors is called *partitional calorimetry* by Winslow, Herrington and Gagge.* By altering the temperature of the air and of the walls of an experimental chamber independently of each other, they were able to study the parts played by convection and radiation under varying environmental conditions.

It is usual to express rates of heat transfer between the body and its surroundings in terms of kilocalories per sq. m. of body surface per hour (Cal./sq.m./hr.),† as in the case of B.M.R. Typical figures for the heat production (M), which can be measured by determination of the oxygen usage of the subject are :

Basal rate	40 Cal./sq.m./hr.
Awake, resting	55 ,, ,, ,,
Light work	75 ,, ,, ,,
Walk four mi./hr.	190 ,, ,, ,,
Strenuous exercise	300–600 ,, ,, ,,

Expressing, in the same units, the heat loss by convection (C), Radiation (R) and evaporation (E), we have the general equation for heat balance, *i.e.* for maintenance of uniform body temperature :—

$$M = C + R + E$$

in which it is agreed that R and C shall have positive values when they cool, and negative when they heat the body (as when surroundings are above skin temperature) : E is always positive.

If thermal equilibrium is not reached, and the body temperature rises or falls, the equation of thermal balance must be adjusted by the introduction of a term indicating the deficit or surplus of heat. This new term is called *storage* (S), whatever its sign, the latter being accounted positive when the body is cooling. The equation now runs :

$$M + S = C + R + E$$

Now M can be measured directly by the rate of oxygen consumption, together with a computation of body surface, as explained in Chapter XLI. We can also measure E from the loss in body weight (*v.* p. 1029) by insensible or sensible perspiration. The value of S (for states of thermal imbalance) can be calculated from the change in mean body temperature and the mean specific heat of the body (0.83). Hence the value of C + R can be calculated.

The following will serve as an example :—

A lightly clad man weighing 65 kg., and 178 cm. height, sits still in a cool room for an hour. During that time his respiratory exchange is found to be at the rate of 302 c.c. O_2 and 257 c.c. CO_2 per min. (R.Q. = 0·85). At the end of the hour his body weight had decreased by 42 g., and his mean temperature by 0·1° C. From these data we calculate as follows :—

(1) Body surface (from Fig. 527, p. 794) = 1·79 sq. m.
(2) Oxygen usage = 18·15 l./hr. At R.Q. 0·85, one l. O_2 = 4·881 Cal. (p. 792).

$$\therefore \quad M = \frac{18{\cdot}15 \times 4{\cdot}881}{1{\cdot}79} = 49{\cdot}\ 5 \text{ Cal./sq. m./hr.}$$

(3) Oxygen intake = 18·15 l. = 26 g./hr. and CO_2 = 30 g./hr.
Now, insensible perspiration = loss of body weight − loss of CO_2 + gain of O_2
= 42 — 30 + 26 = 38 g./hr., which, when vapourised at 0·58 Cal./g. gives

$$E = \frac{38 \times 0{\cdot}58}{1{\cdot}79} = 12{\cdot}3 \text{ Cal./sq. m./hr.}$$

---

* Winslow, Herrington and Gagge. *Amer. J. Physiol.*, 1926, **116**, 641.

† 1 Cal. = roughly 4 B.Th.U. or 0·00004 Therms. It has been proposed to introduce a new unit called a *met*, which = 50 Cal./sq. m./hr.

(4) Body Sp. ht. = 0·83. Cooling = 0·1° C.

$$S = \frac{0{\cdot}83 \times 0{\cdot}1 \times 65}{1{\cdot}79} = 3 \text{ Cal./sq. m./hr.}$$

(5) From the equation

$$\underset{(49{\cdot}5)}{M} + \underset{(3)}{S} = \underset{(40{\cdot}2)}{(C + R)} + \underset{(12{\cdot}3)}{E}$$

we infer that loss by radiation and convection together $(C + R)$ = 40·2 Cal./sq. m./hr.

In further analysis, R can be determined directly by radiometric observation, and so C deduced. Or, alternatively, R and C can be calculated approximately from a knowledge of environmental conditions, skin and clothing temperatures and constants.

When the body is outdoors the problem is greatly complicated by the effects of wind, sun, cloud, etc. For example, surfaces normal to the June mid-day sun in London receive incident radiation at the rate of about 780 Cal./sq.m./hr.

The main regulation of heat loss thus takes place by radiation, conduction and evaporation from the skin. Besides these channels of heat loss, others play a minor part. Thus heat is lost to the body in warming the food and air taken in, and in the setting free of CO_2 from the blood in the lungs, but these are relatively small sources of loss. The following balance sheet represents an example of possible daily losses in an adult man by these different ways :—

EXAMPLE OF HEAT BALANCE SHEET

HEAT PRODUCTION	Cal.	HEAT LOSS	Cal.	P.c. of Total	
Respiratory exchange for twenty-four hours = 497·2 l. O_2 used and 397·8 l. CO_2 produced. R.Q. = 0·8, at which the calorific value of O_2 = 4·827 Cal./l. Hence heat production = 497·2 × 4·827 = approx. . .	2,400	A. *Insensible water loss :* (i) *By breath.* Assume insp. air half sat. at 15° C. (= 6·4 g./cu.m.) and that exp. air saturated at 37° (= 45 g/cu.m.). Then approx. 440g. water is evaporated from resp. passages : at 0·58 Cal./g. . . =	255	11	25
		(ii) *By skin.* Assume loss of 580g. insensible loss at 0·58 Cal./g. =	336	14	
		B. *By Radiation,* say . =	890	37	66
		C. *By Convection,* say . =	711	29	
		D. *Warming Food and Air.* (i) 1,500g. water at 15° warmed to 37°, *i.e.*, by 22° . . =	33	1·5	5·0
		(ii) 1,500g. food at 25° warmed to 37°, *i.e.* by 12° (Sp. ht. = 0·8) . =	14	0·5	
		(iii) 11,500 l. air (= 14,100g.) insp. at 15° and warmed to 37° (Sp. ht. = 0·2374) . =	75	3·0	
		E. *Evaporation of CO_2.* Liberation of 398 l. (= 781g.) CO_2 at 0·108 Cal./ g. . . =	86	4·0	
	2,400	Total . .	2,400	100	

We may say that, under such conditions, about 5 per cent. of the total heat loss of a man not sweating takes place by warming the food and air, 4 per cent. by evaporation of CO_2 from the lungs, about 25 per cent. by the evaporation of insensible water, and about 66 per cent. by radiation and convection. The proportion represented by evaporation of water will, in dry air, increase very largely with high external temperatures, or when heat production is excessive in consequence of violent muscular exercise, because in these cases sensible perspiration will occur.

The relative importance of conduction + radiation and of evaporation, and the extent of storage at different operative temperatures for nude and clothed subjects are illustrated in Fig. 634. In this figure, M and S are given their proper sign: (R + C) and E are, for convenience of plotting, reversed in sign. A few measurements will show that at any ordinate M + S = (R + C) + E, as the equation of heat balance demands.

Physiological Factors. The simple physical picture outlined above is modified by phenomena of a purely physiological nature. These are the result of primary adaptative reactions, *e.g.* changes in the vascularity of the skin, the onset of sweating, or alteration in the water distribution in the tissues. There are also various secondary adaptative responses.

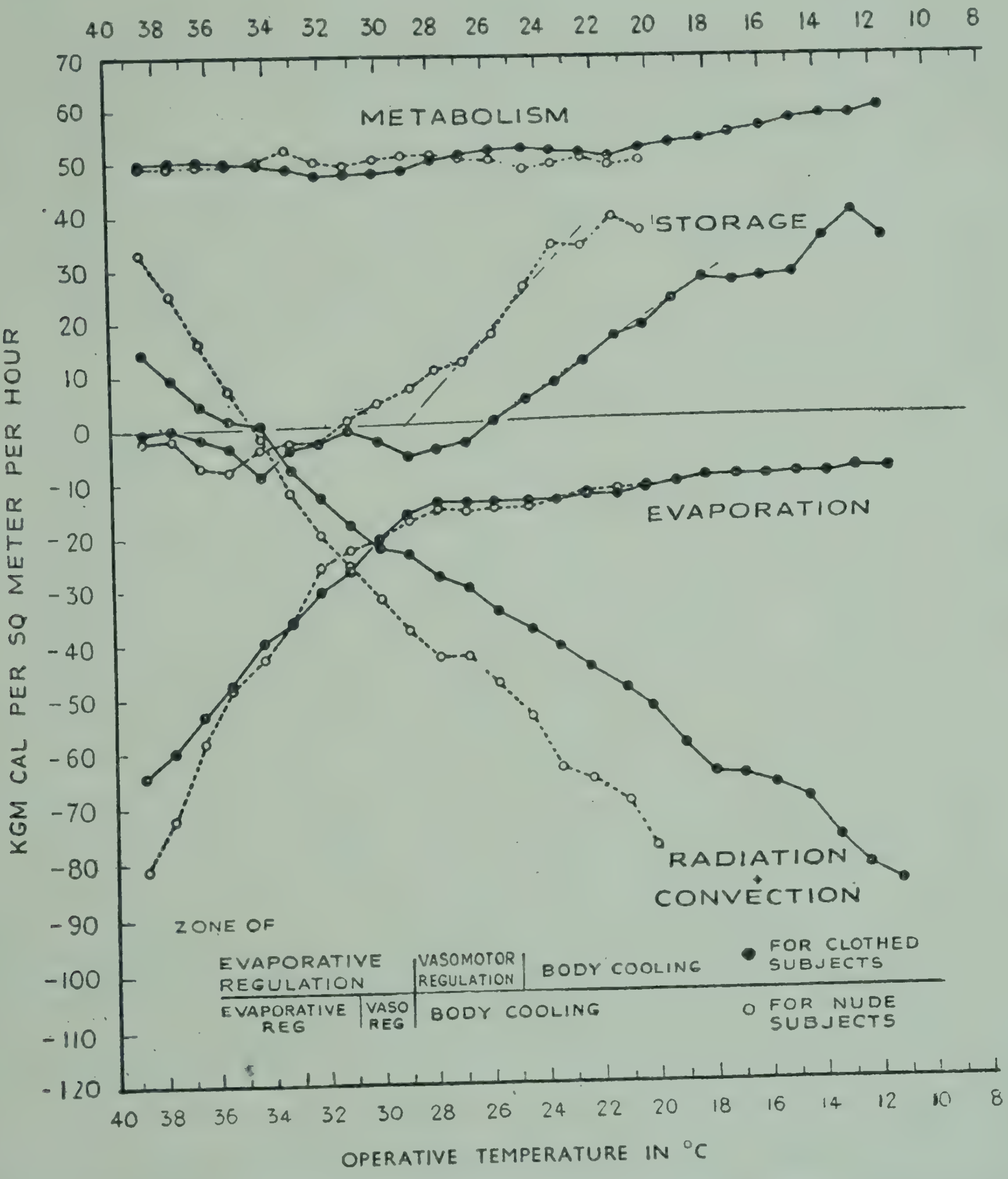


FIG. 634. Thermal interchange for nude and clothed subjects at varying operative temperatures. (GAGGE, HERRINGTON and WINSLOW. *Amer. J. Physiol.*, 1938, **124,** 30.)

At any given level of heat production there are environmental conditions which are nearest to the physiologically neutral, in that they allow of easy and delicate regulation of the body temperature. The further the departure from these conditions, the more physiological factors are called into play: beyond certain limits these prove inadequate, and the body loses or gains in temperature. Thus, resting and lightly clad subjects at a metabolic level of 40–50 Cal./sq.m./hr. will have easily stabilised temperatures when in still air at about 27° C.: if the air temperature rises above that, the skin vessels can be flushed, so that more heat is lost, and if it falls, they can be constricted, so that heat is conserved. This *zone of vasomotor regulation* extends only from about 25°

to 29° C.* Below 25° body cooling begins : above 29° convection can cool the body very little, and at that temperature sweating sets in, and increases in extent and intensity until the maximum is reached : above 35°, all the body heat loss is by evaporation. At a given temperature, which varies widely according to the relative humidity of the air, even evaporative cooling fails to cope with the situation (which may be further worsened by the fact that radiation and convection are now tending to *warm* the body), and the body temperature then begins to rise. We thus have four possible results at progressively increasing surrounding temperatures, viz. body *cooling*, *vasomotor regulation*, *evaporative regulation*, and body *warming*.

Since the rate of loss of heat depends greatly on the difference of temperature between the body surface and the surroundings, and since the temperature of the skin depends on the thermal conductance of the superficial tissues, it follows that a rise of that conductance, due to a diversion of blood to the surface will greatly facilitate the heat loss from the body, and *vice versâ*. In surroundings at 34° C., it has been shown that as much as 12 per cent. of the total cardiac output may be sent to the skin. The supply of blood to the skin of a given area may, in fact, vary over a fifty-fold range, and the conductance of the superficial tissues over a five-fold range (*e.g.* from 11 to 58 Cal./sq.m./hr./° C.), with the consequence that in hot surroundings, or when large amounts of heat are being produced by doing work, the skin temperature *may* be as much as 3° C. above that in temperate surroundings.

These changes in skin vascularity are under the control of the central nervous system, through the vasoconstrictor and vasodilator nerves.† In cold weather, or when the heat production of the body is low, the skin vessels are constricted, the skin is cold, and the heat loss is reduced. The largest fluctuations in vascularity are in the extremities ; here, in the cold, although at first there is vasoconstriction, excessive skin cooling, with risk of ensuing damage to the tissues, is largely prevented by the opening up, by an axon reflex, of arterio-venous anastomoses. In this way the skin surface is flushed with blood, and its cooling delayed. Hence, after handling snow the skin becomes red, warm and hyperalgesic.‡ The anastomoses are controlled, probably centrally, by constriction or dilatation of the arterioles.§ The reaction is closely related to the hyperalgesia in which the " nocifensor " nerves play a part, and also to the spreading flare seen on local damage to the skin.

Plethysmographic measurements of the blood flow through the hand show that the rate of blood flow may range from 0·22 c.c. to 28·3 c.c. per 100 c.c. tissue per minute with change in ambient temperature from 15° to 38° C.‖ In skin it may reach 75 c.c. (finger) ; in muscle it is less per 100 c.c., viz. an increase of 5–10 c.c., but owing to the large bulk of muscle the increase for the whole body may be about 1·5 l/min.¶

At very high external temperatures, vasoconstriction may reappear, and is then useful, since it now, by reducing the skin conductance, retards *intake* of heat by the body.

Evaporative cooling, which undergoes a rapid increase when, at a certain temperature, insensible water loss gives way to sweating, is the method by which heat loss can be effected after other methods have become inadequate. It is well known that the body temperature can remain almost unchanged even in surroundings much above 37° C. Thus Blagdon,** in 1775, showed that a man " could sustain, for a period of fifteen minutes, an exposure to a dry

* These and many other figures apply to persons acclimatised to the conditions of life in the U.S.A. They may be slightly different for European conditions.

† PICKERING. *Heart*, 1932, **16**, 115.

‡ LEWIS. *Heart*, 1930, **15**, 177, 351.

§ GRANT *et al.* *Heart*, 1931, **15**, 385 ; 1932, **16**, 69.

‖ FORSTER *et al.* *Amer. J. Physiol.*, 1946, **146**, 600.

¶ BARCROFT *et al.* *J. Physiol.*, 1947, **106**, 271.

** BLAGDON. *Phil. Trans.*, 1775, **65**, 111, 484.

atmosphere of 250° F. without ill-effects, while a beefsteak exposed to the same temperature was cooked in thirteen minutes."

This is because there is a profuse secretion, and evaporation, of sweat over a large surface of the body. The critical air temperature for sweating is, for nude subjects, 31° C., and for lightly clad ones 29° C., the *average* skin temperature being 34°: sweating increases sharply with skin temperatures of 35° or over. Since the evaporation of 1 g. of water absorbs 0·58 Cal., the evaporation of 120 g. of sweat would lower the body temperature of a 70 kg. man at least 1° C.; this often results in the skin temperature being below that of the surrounding atmosphere. The skin temperature of a man in hot surroundings may for this reason be lower when he is doing work than when at rest: this, by reducing his loss by R + C would be a slight disadvantage.

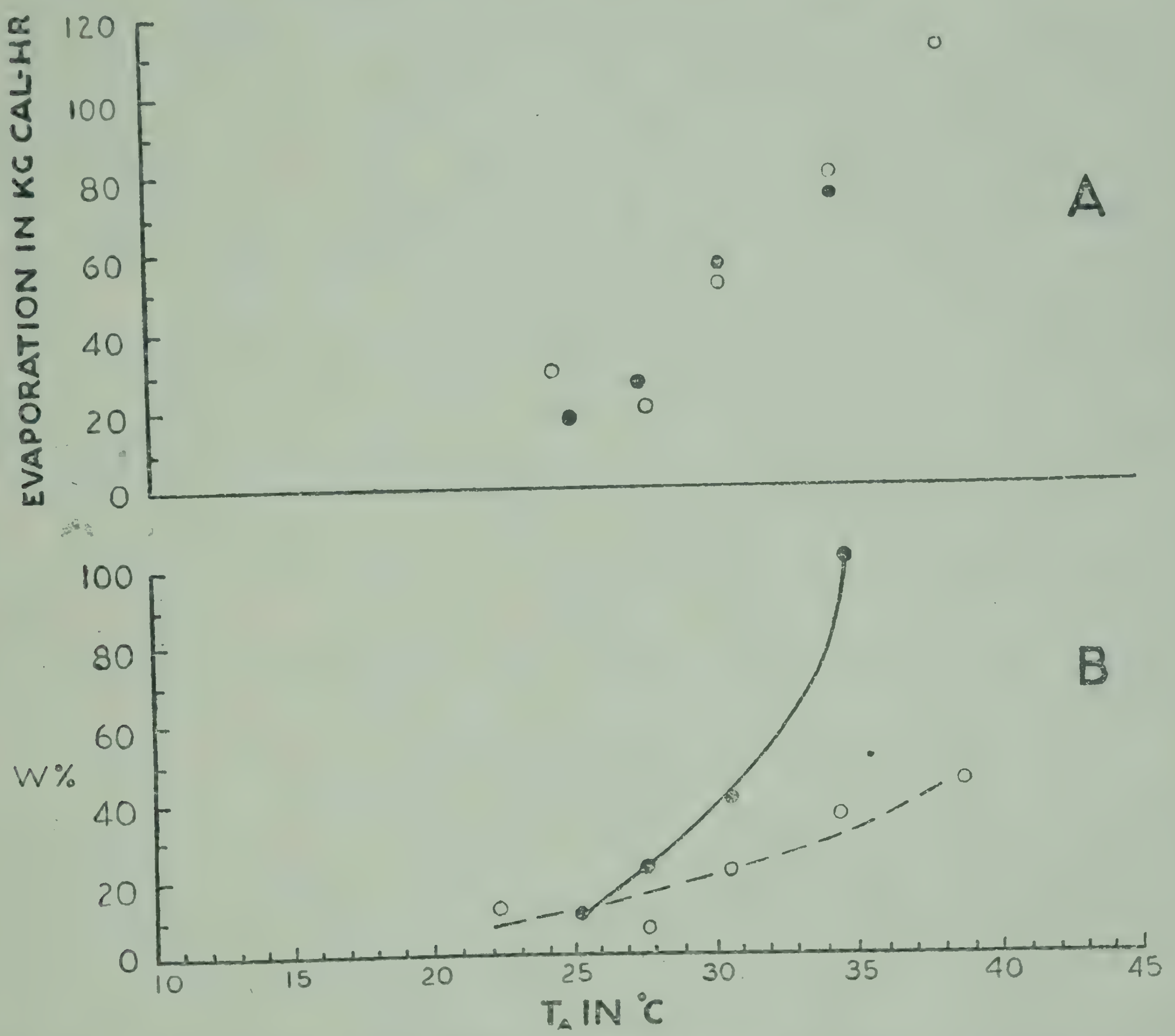


FIG. 635. Evaporative heat loss (A) and per cent. of maximal possible area of wetted skin (B), at varying air temperatures: nude subjects. Solid circles = high humidity (40–80 per cent.): open circles = low humidity (14–40 per cent.). (After WINSLOW, HERRINGTON and GAGGE. *Amer. J. Physiol.*, 1937, **120**, 288.)

If, as happens in rare cases, a man has no sweat glands,[*] or if evaporation be impeded by placing him in a hot atmosphere saturated with aqueous vapour, the temperature of the body runs up rapidly, and death may ensue from hyperpyrexia or heat-stroke. Although a man can stand exposure to a temperature of 120° C. (250° F.) for a short period if the air is dry, a temperature of 32° C. (90° F.) may in time be fatal if the air be saturated with moisture. Possibly the reports that death may result from entirely coating the skin with an impervious layer may be due to a similar cause. Cases are on record of boys dying after being varnished and gilded in order to feature in carnivals, *e.g.* one in Rome in 1473, and one in Florence in 1513, with fatal results.

Sweating also comes into play when the heat production of the body is very largely increased, as by violent muscular exercise, and a man may then sweat profusely at air temperatures below freezing point.

---

[*] RICHARDSON. *J. Biol. Chem.*, **1926**, **67**, **397**.

The vasodilatation of the skin which accompanies sweating also assists in cooling by bringing heat to the skin, and by providing the necessary water for formation of the sweat.

The value of $W\mu$, which is a measure of the wetted area of the skin, is small and constant below 30° C., because the evaporation is all insensible; above this, its value at a given temperature and air speed increases with increase in the humidity of the air, until it reaches a maximum, *i.e.* as the evaporative power of the air is reduced, there is a compensatory increase in the wetted area, so that the total evaporative heat loss tends to remain constant (Fig. 635). Or alternatively, we may say that the relative humidity of the skin increases with the humidity of the air, and this again keeps evaporative loss constant.

The maximum evaporative cooling, reached when the maximum area is wetted, is about 30 Cal./sq.m./hr./cm.Hg difference of water vapour pressure $(P_s - P_a)$ and is reached at about 43° C. nude, or 52° C. clad, with low humidity. With high R.H. of air, the clad subjects adapt less well, and temperature regulation fails at T_o 35° C. or lower.

Shifts of fluid within the body also play a part in the heat regulation of the body: the best indicator of fluid shift from cell interspaces to plasma is the concentration of serum protein.* This shows a fall soon after exposure to heat, followed by a rise if sweat loss is not made good. The effect of the initial dilution of the blood is to keep the vascular system filled well, despite vasodilatation and deflection of blood to the skin, and to augment the thermal conductance of the surface tissues. Conversely, exposure to cold at first leads to passage of "water" from blood to tissues; later, due to chilling of the nervous centres, the trend is reversed, and the excess of extracellular fluid may lead to œdema.†

THE EFFECT OF CLOTHING

Clothing protects from high, as well as from low, environmental temperatures. It protects in proportion to its low conductivity for heat (*e.g.* linen 0·00021, silk 0·000095) so that there is a steep temperature gradient from skin surface to clothing surface, from which the heat is dissipated by R and C to the surroundings, at a rate depending on their respective temperatures. Since the thermal conductivity of air (0·00005) is even lower than that of most clothing materials, while its specific heat is also low (0·24 cal./g.), it is evident that the air beneath and between layers of clothing, and within its pores, has great value as a thermal insulator. This explains why two layers of thin clothing are better than one layer of double the weight. Solid and impervious materials are, for a similar reason, poor in insulating power, and are avoided as clothing materials, unless they have some special advantage, such as that of being waterproof: they have the further drawback of not allowing of ready evaporation from the skin. Clothing increases the effective radiating surface by about 16 per cent., but lowers the temperature of radiating surfaces more than proportionally, and so cuts down radiation loss indoors by about 40 per cent. (Fig. 634).

The distribution of temperature on the skin of the clothed body is different from that in the nude, the skin of the extremities being cooler, and that of the trunk warmer, than in the nude. But the mean skin temperatures are about the same in the two states, when the rates of heat loss by R + C are equal.

It is evident that, if the body temperature is to remain constant, all the heat passing through the skin must, where clothing covers it, also pass through the clothing. When sweating begins, skin temperatures and conductances are higher for clothed than for nude, and evaporation of sweat mainly occurs at the skin surfaces, and not from folds of the clothing: hence the clothed body loses heat by R + C, and in general responds in much the same way as the nude, though at a different level. When at rest, body cooling sets in at $T_o =$ 8° C. below that of the skin, as compared with 4° below for the nude.

---

* Lee and Mulder. *J. Physiol.*, 1935, **84**, 279, 410.

† Barbour *et al.* *Amer. J. Physiol.*, 1943, **140**, 9.

Thus, for M − E = 40 Cal./sq.m./hr., and at $T_o = 25{\cdot}1°$ C., the gradients for light clothing would be :—

Skin temp. .	33·1°	T. diff. = 2·5°. Physical conductance of clothing = 15·8 Cal./sq.m./hr./° C.
Clothing temp. .	30·6°	T. diff. = 5·5°. Actual conductance for loss by R + C 7·3 Cal./sq.m./hr./° C.
Operative temp. .	25·1°	

The physical conductance of the clothing is 40/2·5 = 15·8 units per ° C. diff. in temperature between the skin and clothing temperatures, while the actual conductance by R + C from the surface is 40/5·5 = 7·3 units. We may take it that ordinary light clothing at indoor temperatures at rest transmits about 38 out of 50 Cal./sq.m./hr., which = 76 per cent. of the heat production, for subsequent dispersal by R + C. Under common conditions indoors, for clothed persons at rest, about 45 per cent. of the total loss by R + C is effected by convection : for the nude it would be about 52 per cent., *i.e.* clothing chiefly reduces the relative convection losses. At temperatures below 30° C. skin conductance is lower for nude than for clad : above that the conductance rapidly increases, especially for the clad.

ASSESSMENT OF THE EFFECTS OF ENVIRONMENTAL CONDITIONS

The physiological properties of a given environment depend essentially upon the cooling (or warming) power of the surroundings upon the skin. This varies over an extremely wide range, since it depends upon differences of temperature, humidity, and rate of movement of the air. There is further the influence of the interchange of radiant heat between skin or clothing and surroundings.

The conditions indoors and outdoors, are, under ordinary circumstances, fundamentally different, the cooling and evaporating powers being greatly reduced indoors, owing to the stillness of the air.

Air temperature is obtained by use of an ordinary thermometer, but this gives only a limited indication of the physiological aspects of the atmospheric conditions.

Air humidity is given by use of the wet- and dry-bulb thermometer (or, better, by the whirling psychrometer); this gives more information, because the evaporative power of the atmosphere is indicated by the depression of the wet-bulb instrument. The average of the percentage humidity and dry bulb temperature ("humature" scale) has been used as a rough guide to tolerability.

Radiation can be assessed most simply by the globe thermometer.* This is a hollow 6-in. metal sphere, painted matt black, and having an ordinary thermometer with its bulb at the centre of the sphere. The reading of this thermometer will be higher or lower than air temperature according to whether the surroundings are at higher or lower temperatures than the air.

Air velocity can be measured by the anemometer, or by the rate at which a warm silvered surface (silvered Kata thermometer) cools down.

Dufton's scale of *equivalent temperatures* takes into account the air temperature, air movement, and radiation. It is defined as that temperature of an uniform environment in which, in still air, the body would lose heat at the same rate as in the environment in question. It can be measured by the *Eupathioscope*, or comfort indicator, which is a heated blackened cylinder 22 inches high by 7·5 inches diameter : its heat input, electrically run, is so regulated that, in the particular environment under study, the surface temperature of the instrument is that of the mean surface of the clothed body. Th equivalent temperature can also be estimated from readings of the globe thermometer (radiation), the kata-thermometer, and the ordinary thermometer, and reference to tables or nomograms.

---

* Bedford and Warner. *J. Hyg.*, 1934, **34**, 458.

The use of *operative temperatures*, and of partitional calorimetry, is probabl the most scientific method of approach in studying the thermal relationshi of the body to any given environment, since it enables the problem to b regarded chiefly as a physical one.

Effective Temperature. Another method of assessing climatic conditions * is by the estimation of what is called the effective temperature of Yaglou. This has been empirically devised by direct observation on human subjects and takes into account temperature, humidity radiation and movement of the air.

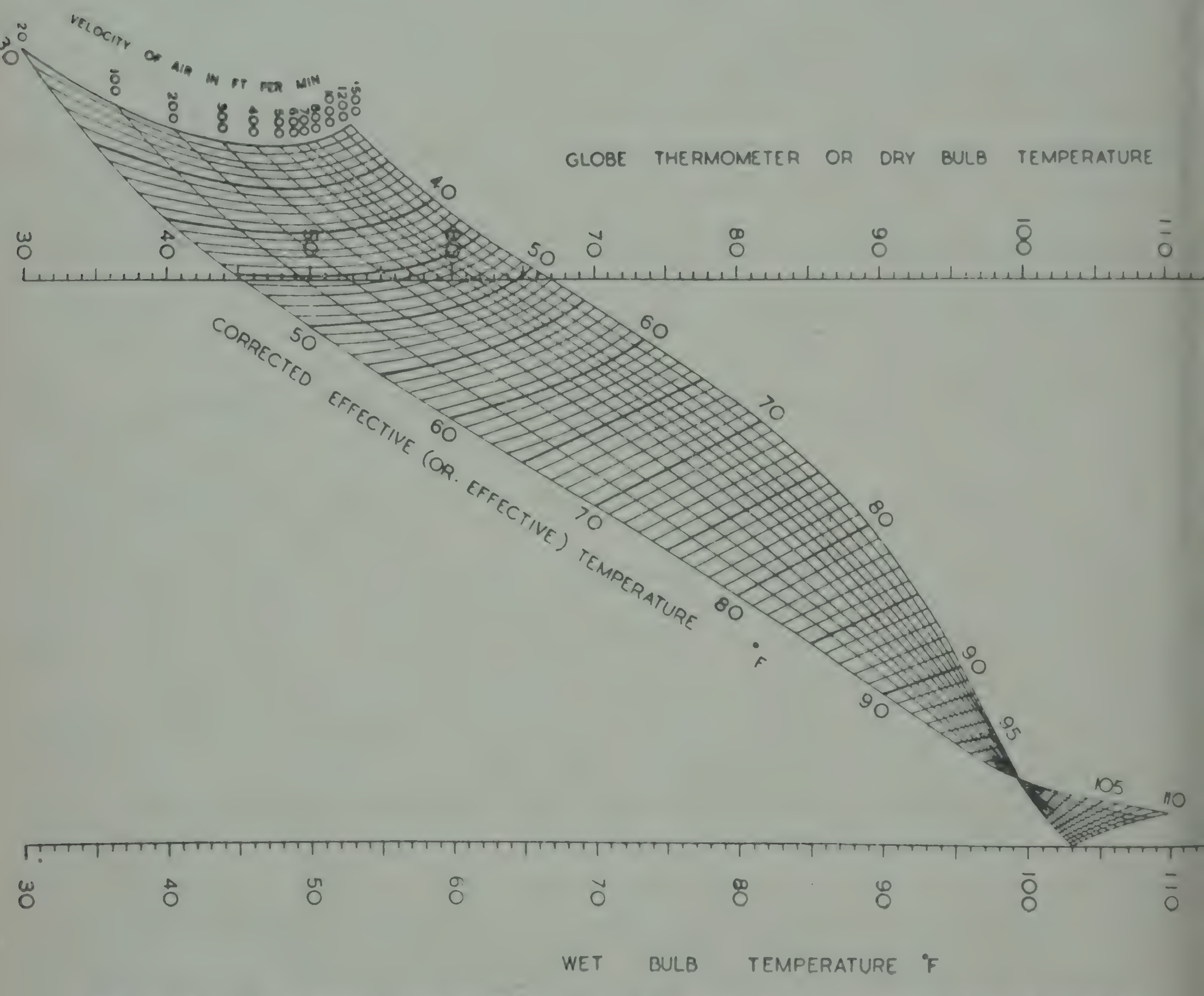


FIG. 636. Nomogram for determination of effective temperature on men normally clad and doing light work. (Scales in deg. F.) (BEDFORD. M.R.C. *War Memo.* No. 17, 1946.)

The effective temperature of an atmosphere is the temperature of a still, saturated atmosphere which would produce the same feeling of temperature as the atmosphere in question. It is thus a practical comfort criterion derived from subjective experience. It is found by determining globe or dry-, and the wet-bulb temperatures and average air velocity. Then by reference to a nomogram (Fig. 636) the effective temperature is read off directly. The result is called the corrected effective temperature when the globe thermometer replaces the dry-bulb air thermometer, because it allows for radiation effects. Thus a globe thermometer temperature of 76° F. (24·5° C.) and a wet bulb of 62° F. (16·5° C.), with air velocity of 100 feet per minute, gives, a corrected effective temperature of 69° F.

---

* LEE and COURTICE. *Trans. R. S. Trop. Med. Hyg.*, 1940, **33**, 601; BEDFORD, M.R.C. *War Memo.* No. 17, 1946 (H.M.S.O.).

Climate and Comfort. The important physiological effects of climates were recognised by Hippocrates, but still present many problems. The tolerability of a climate depends on the balance of the effects upon the body of temperature, relative humidity, radiation, and air movement : the effects are modified by clothing, and by physical exertion.

Extremes of temperature or humidity are usually unpleasant, as also is a monotonously uniform climate. Probably ideal conditions for light exertion would be at a temperature of 19·5° C. (67° F.) and relative humidity 50 per cent. : countries at present occupied by races that regard themselves as the most energetic and civilised have variable climates, with rather low mean temperatures, though man is very adaptable in these respects.

It is possible to calculate the limiting conditions for heat stroke for an individual under given conditions of climate if we have certain data, such as the value of M, the metabolic rate : v, the air velocity in m./sec. ; T_s and T_w, the temperatures of skin and surroundings, and I_s and I'_d the total heat content of air saturated at temperature T_s, and at the wet-bulb temperature respectively. At " rest " M is normally about 47. If work is being done at the rate of 1,000 × W. ft./lbs. per minute, then, approximately, $M = 64 + 35W$, Cal./sq.m./hr.

The curves showing limits beyond which control of body temperature is lost are derived from the use of the formula $M = 3{\cdot}6\,(T_s - T_w) + \dfrac{200}{3}\, v^{\frac{1}{2}}\,(I_s - I'_d)$, and some are given in Fig. 637.

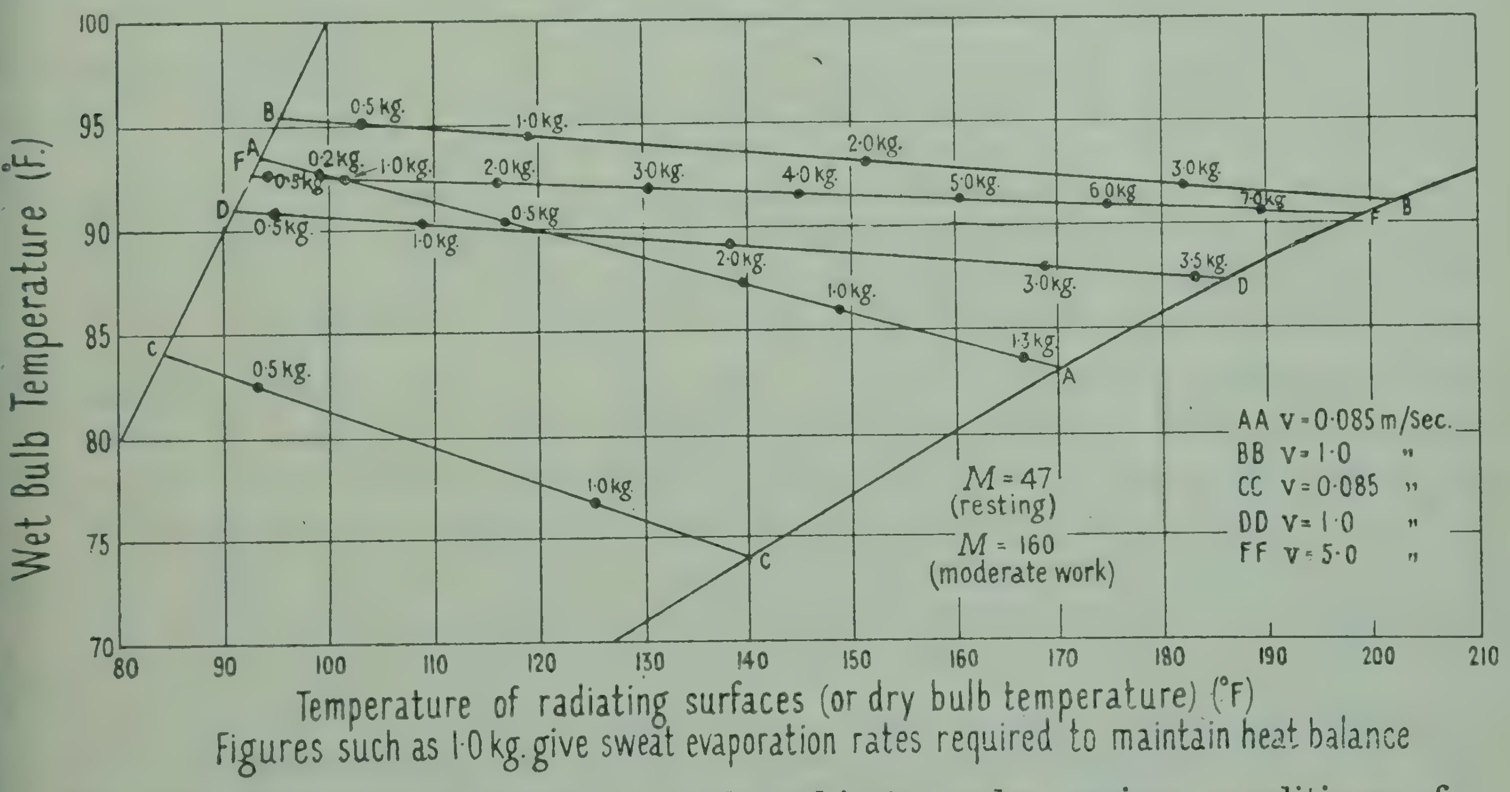


FIG. 637. Heat-stroke limits for nude subjects under various conditions of physical activity and air-movement as shown in figure. A and B at rest. C, D and F at work. Above the lines heat stroke is liable to occur. Figures such as 2·0 kg. show sweat per hour needed to maintain heat balance. (BRUNT. *Proc. Physical. Soc.*, 1947, **59**, 713.)

Thus curve AA shows that men resting in nearly still saturated air cannot tolerate a temperature over 93·5° F. The temperature in the Black Hole of Calcutta has been said to be 34° C. (93·2° F.). Comparison of curves C and F shows the large effect of increasing air movement from 0·085 m./sec. to 5 m./sec.

In Fig. 638 is shown a tentative classification of climates derived from these curves. The line AA shows the heat-stroke limits for a nude man resting in still air ; BB the same but with air moving at 200 ft./min. ; CC, the limiting conditions for clothed man resting in sunshine, and with about a third of the skin wetted with sweat ; DD, the same, except that he is walking at 3 mi./hr.

To the left of these lines temperature and humidity are such that the body temperature will not rise, while to the right of them there will be a rise of body temperature. Each line is thus a climatic boundary of tolerability for this subject. For example, a temperature of 90° F. at 80 per cent. humidity is risky for a clothed man, although 100° F. at 20 per cent. humidity could be borne, though, of course, it would be "very hot." The "rawness" of damp, and the "keenness" of dry, air in cool climates are also illustrated in the figure.

The reduction in risk from tropical diseases which advance in medical science has wrought now raises the question whether many tropical areas might not now be more densely populated. The study of climate is thus of great importance. Hot climates show two extremes, which differ profoundly in physiological effects, according to whether the air is dry or moist. In hot desert climates the air is dry and clear: radiation conditions are good, and the ground, being dry, heats and cools quickly, so that surface ground temperatures may be 180° F. (82° C.) by day and 60° F. (15·5° C.), or less, at night: the air temperature, therefore, also shows large fluctuations, *e.g.* 113–68° F. (45–20° C.). Relative humidity varies from 5 to 30 per cent., and winds are frequent.

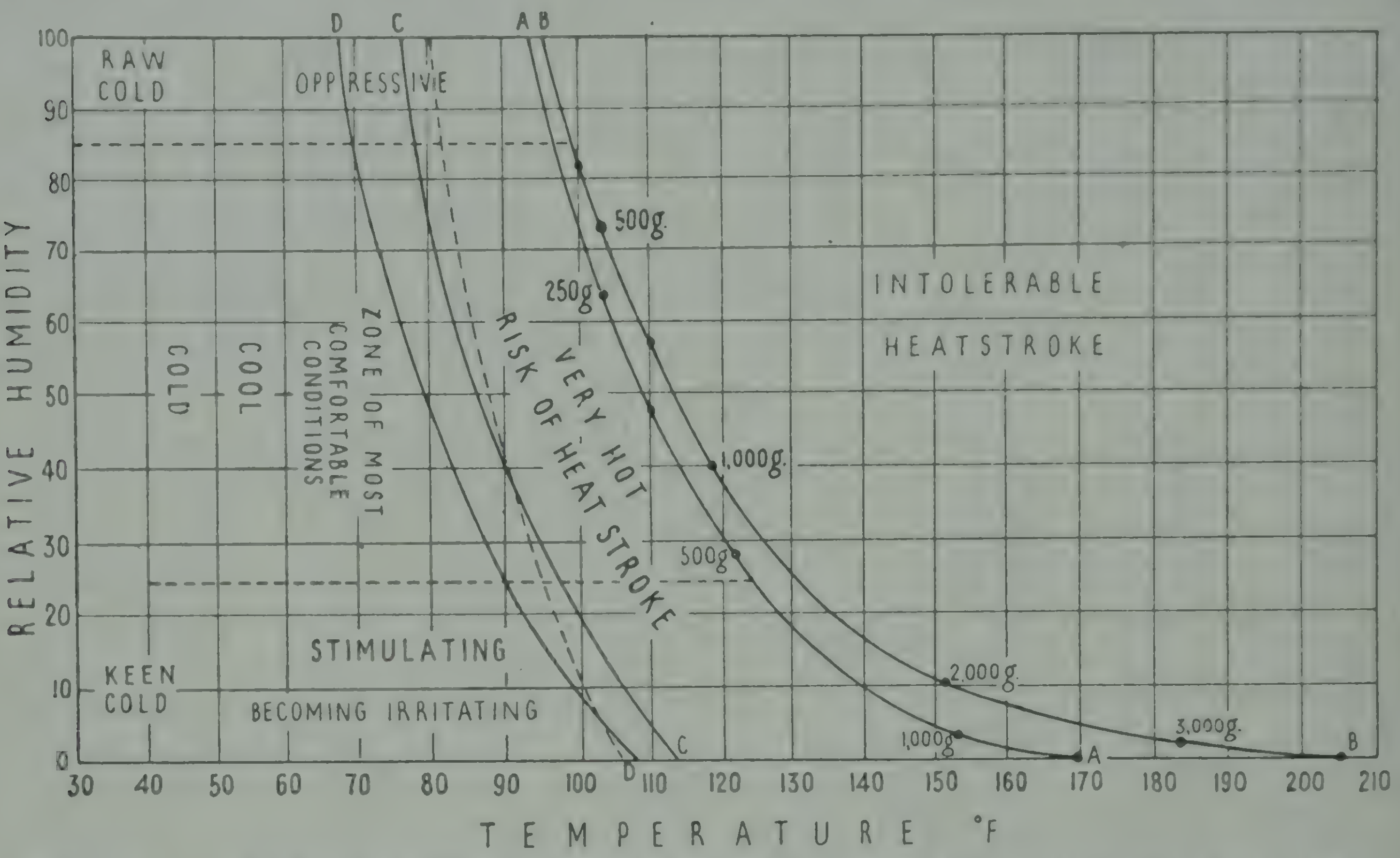


FIG. 638. Tentative classification of climates. For description see text. The broken line = equivalent temp. 80° F. The figures, 500 g., etc., indicate rate of evaporation of sweat in g./hr. necessary to maintain heat balance in man of average size. (After BRUNT. *Proc. Physical. Soc.*, 1947, **59,** 713.)

Wet tropical (jungle) climates differ from the above in all save the high temperatures. These, say 71–92° F. (22°–33° C.), are not often so high as in desert climates, but are likely to remain uniform for long periods, ground and air temperatures in the jungle often varying less than 1° C. in the twenty-four hours. There is frequent and heavy rain, cloudy skies are common, and, in jungles, with their foliage cover, the air is stationary, and, most of the time nearly saturated with moisture (R.H., *i.e.* relative humidity, at least 80 per cent.).

Reference to Fig. 638 will enable us to predict that the jungle climate, say, 85° F. and 85 per cent. R.H., is likely to be much more trying than that of shade in the desert, say 100° F. and 15 per cent. R.H.: light work would be possible in the latter conditions, because, although radiation and convection both *add* heat to the body, the evaporation of sweat would be rapid, especially when there was wind. It must be remembered that in sunshine, however, the body may be heated by radiation at a rate considerably greater than its own metabolism. Artificial cooling and air-conditioning of dwellings in the tropics is mainly a matter for the future.

The unpleasantness of cold, raw days is due to an increase in the conductivity of the clothing owing to rise in its moisture content.

Except for the risk of frost-bite, cold climates, though often less comfortable, are more readily dealt with than hot ones, since heated dwellings, increased clothing and physical exertion can largely compensate for the conditions.

For *comfort*, the skin temperature of resting subjects should be about 33° C., and during work about 31° C., and these would be the conditions when at rest at about 82° F. and 50 per cent. R.H., or during moderate work with slight air movement at 61° F. and 50 per cent. R.H.

For indoor conditions, habituation is an important factor: the Japanese in winter regularly support degrees of cold most uncomfortable for Europeans, and the French shiver in our rooms, as we do in their out-door cafés. In the U.S.A., 72° F. and 50 R.H. is considered ideal for winter, but in Great Britain, as many Americans can now testify, 65° F. is considered warm in winter rooms, and 60° F. with R.H. above 80 per cent. is not unusual. Clothing in the two countries is adjusted to suit these different standards. Probably a summer temperature of 66° F., R.H.50, is ideal for resting subjects indoors.

Acclimatisation to Heat. Study of the adaptation to altered climatic conditions is of greater importance now than ever before, because by air transport a person can in a few hours be translated to a totally unfamiliar climate.

The chief immediate physiological responses to exposure of the body to heat are :—increase of respiratory ventilation, with rise of blood pH, excretion of alkaline urine, increased pulse and circulation rate, sweating, dilatation of arterioles and capillaries in the skin, usually a fall of arterial pressure, an increase of blood volume, rise of body temperature, and in extreme cases rise of basal metabolic rate. If exposure is continued, say for several days, either the climate proves insupportable, or acclimatisation occurs. Sweating then occurs more easily, and the sweat is more dilute. The urine becomes more concentrated and may at times be free from chlorides. Skin temperatures may remain high, but the resting body temperature returns to normal, though exercise may lead to a rapid rise, especially in humid climates.

How the acclimatisation is brought about is obscure, but when it has occurred, temperature regulation is more prompt and efficient. One factor is probably a loss of body weight, whereby a smaller mass of metabolising tissue now exposes a greater cooling surface. Adjustments of the central nervous system, of endocrine activities, of circulation, and of water distribution are other factors. Acclimatisation to work in hot, dry climates, whereby comfort is increased, pulse rate reduced and skin and rectal temperatures lowered, occurs in a matter of a few days, some 80 per cent. of that reached in three weeks being attained in one week *: the acclimatisation once established may last for two or three weeks.†

BREAKDOWN OF HEAT REGULATION

Regulation of the body temperature is only possible if the demands of the environment are not unreasonable. Beyond a certain point in either direction it breaks down, the body temperature alters, and may ultimately reach a point at which death occurs.

The human body, with its artificial aids, can tolerate a greater alteration of the surrounding, or of its own, temperature, towards cold than towards heat. Thus the environmental temperature can fall 55° C. below a temperate level, and the body temperature 12° C. below "normal," without inevitable death: in the upward direction the limits would be about 30° C. above temperate for environment, and 6° C. above "normal" in body temperature.

---

* Robinson *et al.* *Amer. J. Physiol.*, 1943, **140**, 168.

† Henschel *et al.* *Ibid.*, p. 321.

Lowering of body temperature can be effected by removal of clothing, cold baths or ice-packs, general anæsthesia, eating ice, or exposure to cold air. As the temperature falls the nervous system becomes sluggish, consciousness is gradually lost, and, at rectal temperatures of about 25° C. most reflexes have disappeared. Recovery has been known after a fall of body temperature to 18° C.

Increase of body temperature can be brought about by exercise, increase of clothing, hot moist air, hot baths, hot drinks, radiant heat, diathermy, heat puncture of the nervous system, or by the injection of pyrogenic substances, *e.g.* nitro-naphthol, tetrahydro-β-naphthylamine, yeast, bacterial toxins, or in fevers. When the high temperature is maintained, so that adjustment ultimately fails, the effects take one of the following five forms: hyperpyrexia, heat exhaustion, heat cramp, dehydration and thermal anhidrosis.

Hyperpyrexia, or true "heat-stroke," is due to damage to essential tissues. The rectal temperature is above 104° F., and may continue to rise. Various nervous phenomena are seen, *e.g.*, nausea, headache, stupor or convulsions.

Heat exhaustion is a circulatory failure. There is low arterial pressure, fatigue, syncope, often a subnormal temperature, and sweating.

Adequate NaCl ingestion is an important factor in preventing heat exhaustion.*

Heat cramp is due to chloride deprivation and loss through sweating: There are muscular cramps, low Cl in blood serum and urine, normal temperature and pulse, vomiting, polyuria, diarrhœa and loss of weight. It is rectified by administration of salt.

Dehydration may occur alone, without salt loss: there is thirst, reduced salivation, oliguria, normal temperature, shrunken appearance, exhaustion and various nervous phenomena.

Thermal anhidrosis is preceded by profuse sweating (hyperhidrosis): then sweating ceases except on face and neck. There is no Cl loss, hyperpyrexia, or increase of pulse and respiration rate. Collapse and unconsciousness may follow: recovery is rapid when removed to cool surroundings.†

Fever. In fevers, the body temperature is raised for hours or days, although the environment provides no explanation for it. Opinions have differed as to the respective importance, in the production of febrile states, of increased heat production and of diminished heat loss. One thing is certain, that in many cases the heat-regulating mechanism is deranged, and is "set" at a higher level than normal.

It is true that an increased rate of metabolism is found to accompany most febrile states, but it is more likely to be as a result, rather than the cause, of the rise of temperature. In the case of some pyrexias of chemical origin, however, *e.g.* as a result of the injection of pyrogenic toxins, there is some evidence of a primary stimulating effect on the general metabolism, of which the rise of temperature would then be the consequence.

As regards heat loss there is less room for doubt, since it has been shown in a variety of ways, and in divers pyrexial states, that disturbance of heat loss is an important factor. The skin is pale, and at first cooler than usual (the subject shivers), and the circulation rate in the skin has been shown to be reduced. The skin is also dry, in spite of the ultimately high body temperature, so that evaporative cooling is in abeyance. The rising temperature, together with the shivering, creates a vicious circle by increasing the heat production. Sooner or later the body temperature rises to such a level that, in spite of the steep gradient to the skin, the latter has got hot enough for loss by R + C to be adequate to balance the heat production, and at this high level the body temperature may remain fairly constant, though with exaggerated diurnal fluctuations.

Usually, at the onset of fever, there is definite loss of water from the blood:

---

* Taylor *et al.* *Ibid.*, *p.* 439.

† Wolkin *et al.* *J. A. M. A.*, 1944, **124**, 478.

this has been ascribed to an increased binding of water by the other tissues, such as muscle. Its effect is to reduce the blood volume, favouring its withdrawal from the skin to the internal organs and thus aiding in the restriction of heat loss. In fact, the importance of anhydræmia may be very great, as is shown by the fact that pyrexia can be produced by the mere abstraction of water from the blood as a result of intraperitoneal injection of hypertonic salt solutions.

As the fever passes off, quickly or slowly (crisis or lysis), these causes cease to operate : sweating occurs from a flushed skin, water is recalled from the tissues to the blood, the vicious circle is broken, and the temperature falls to normal.

Antipyretic drugs, such as aspirin, act by lowering the setting of the deranged heat regulating centres to the normal level, and mainly increase heat loss by sweating, vasodilatation and hydræmia.*

THE NERVOUS REGULATION OF TEMPERATURE

The accurate balance between heat production and heat loss involves co-ordination by the higher levels of the central nervous system. A spinal, or a "decerebrate" animal cannot regulate its body temperature, but an animal merely deprived of its cerebral cortex can do so ; so also can one deprived of all brain substance anterior to the middle of the optic thalamus.†

There is good evidence that the important function of integrating the various mechanisms concerned with temperature regulation is located in the hypothalamus, and in man is probably located in the substantia grisea, near the floor of the third ventricle.‡ Both heat production and heat loss are there controlled, and there are separate centres for response to heat and to cold. That for heat is situated in the anterior hypothalamus, and that for response to cold in the caudal part of the lateral hypothalamus.

If the anterior part of the hypothalamus is warmed, the processes causing heat loss, *e.g.* vasodilatation, panting, sweating, etc., are brought into action § ; this suggests that some nuclei here act as thermoregulators, responding directly to a change in temperature of the blood. Electrical stimulation of the hypothalamus may also produce similar responses, while lesions of the anterior hypothalamus lead to a loss of power to withstand high temperatures.‖ It may be concluded therefore that the general response to rise of temperature is directed from the anterior hypothalamus.

Lesions of the posterior part of the hypothalamus tend to cause subnormal body temperatures ¶ ; animals so operated exhibit vasodilatation, and fail to shiver as the temperature drops. It is inferred, therefore, that responses to falling temperature are controlled mainly by the posterior or sympathetic group of hypothalamic nuclei.

The lower centres for heat loss must be placed in the medulla and cord, at any rate so far as concerns control of heat loss by alterations in the blood supply to the skin, or by the secretion of sweat. The same is true for the heat production arrangements, such as shivering, vaso-constriction, metabolic changes, etc. But these lower centres can only function in a co-ordinated manner when under the higher control of the hypothalamus.

It is evident that the heat-regulating centres are put into action as a result of a change in the temperature of the blood reaching the brain : but it is probable that they are also reflexly affected, as *e.g.* by the application of heat or cold to the skin. If one leg is placed in hot water, a reflex vasodilatation occurs in the other limb, even if the return of heated blood from the warmed leg is prevented by occlusion of its circulation.** Local sweating of a reflex nature can also be caused by application of heat to the skin.

---

* GUERRA and BARBOUR. *J. Pharmacol.*, 1943, **79**, 55.

† ISENSCHMID and KREHL. *Arch. f. Exp. Path. u. Pharm.*, 1912, **70**, 109.

‡ BAZETT *et al.* *Arch. Neurol. and Psych.*, 1933, **30**, 728; FRAZIER *et al.* *Brain*, 1936, **59**, 122.

§ MAGOUN *et al.* *J. Neurophysiol.*, 1938, **1**, 101.

‖ RANSON, FISHER and INGRAM. *Arch. Neurol. Psych.*, 1937, **38**, 445.

¶ CLARK, MAGOUN and RANSON. *J. Neurophysiol.*, 1939, **2**, 61.

** DUTHIE and MACKAY. *Brain*, 1940, **63**, 295.

The reflex vascular and other responses to temperature changes at the surface are probably more delicate than those due to direct effects of altered temperature of the blood on nervous centres. This is especially so with responses to cold, the skin receptors for which are more numerous and more superficial than are those for heat: the threshold difference for sensations of cold is also very small, viz. a drop at the rate of 0·004° C./sec. The skin of the forehead is very sensitive and can detect a rise of 0·003° C. in three seconds.

Shivering is a reflex phenomenon, though it may also be of central origin, and due to fall of blood temperature. A fall of temperature of 0·5° C. to 0·7° C. almost always results in shivering, together with a feeling of internal cold. Administration of alcohol can postpone the onset of shivering until the mean body temperature has fallen by 1° C. or more; taking ice, or iced water, advances it, so that it may occur with a lowering of only 0·2° C. The application of a cold stimulus to an area of skin may, however, cause shivering, although the change of body temperature may be almost inappreciable, or even if the body temperature is rising as a result of heating another area. It would appear, therefore, that sensory stimulation by cooling the skin is a more delicate and potent cause of shivering than is a fall of general body temperature. The cooling of an area of insensitive skin, *e.g.* in a subject with divided spinal cord, does not cause shivering unless there is a definite fall in body temperature as well.*

It has been shown that the reflex of shivering is controlled, not from the spinal cord, but somewhere anterior to the mid-brain,† and almost certainly in the posterior hypothalamus. The efferent paths are not the pyramidal tracts, since damage to them does not affect it, but some tracts in the anterolateral columns, probably the tectospinal or rubrospinal tracts.‡ Normal shivering depends on the integrity of cerebellar function, and on the presence of intact afferents from the muscles.

When shivering occurs, the heat production may be more than doubled; but the heat loss is also increased somewhat, because the increased air movement raises the loss by convection.

For Reference

ADOLPH (1947). Physiology of Man in the Desert. New York and London: *Interscience Publ.*

BAZETT (1927). Physiological Responses to Heat. *Physiol. Rev.*, **7**, 531.

BEDFORD (1946). Environmental Warmth and its Measurement. (M.R.C. *War Memo.* 17. H.M.S.O.)

BRUNT (1943). The Reactions of the Human Body to its Physical Environment. *Q. J. Roy. Meteor. Soc.*, **69**, 77.

BRUNT (1947). *Proc. Physical. Soc.*, **59**, 713.

DEIGHTON (1933). Physical Factors in Body Temperature Maintenance and Heat Stimulation. *Physiol. Rev.*, **13**, 427.

DILL (1938). Life, Heat and Altitude. (Oxford Univ. Press.)

DU BOIS (1937). The Mechanism of Heat Loss and Temperature Regulation. (Lane Med. Lects. Stanford Univ. Press.)

DU BOIS (1938). Heat Loss from the Human Body. (*Harvey Lects.*, 1938–39, p. 88.)

GIAJA (1938). Homéothermie et Thermorégulation. (Actualités Scientif. et Industr.) (Paris: Hermann.)

LEE (1940). A Basis for the Study of Man's Reaction to Tropical Climates. (Univ. Queensland Papers.)

HARDY (1937). Physical Laws of Heat Loss from the Body. *Proc. Nat. Acad. Sci.*, **23**, 631.

MACHLE and HATCH (1947). Heat: Man's Exchanges and Physiological Responses. *Physiol. Rev.*, **27**, 200.

PEMBREY (1898). *Art. Animal Heat*, Schäfer's Text-book of Physiology (I.), 785. (Edin. and Lond.: Pentland.)

SUNDSTROEM (1927). Physiological Effects of Tropical Climate. *Phys. Rev.*, **7**, 320.

---

* JUNG, DOUPE and CARMICHAEL. *Brain*, 1937, **60**, 28.

† SHERRINGTON. *J. Physiol.*, 1924, **58**, 405.

‡ UPRUS, GAYLOR and CARMICHAEL. *Brain*, 1935, **58**, 220.

BOOK VIII

THE PROVISION OF SPECIAL CHEMICAL STIMULANTS, AND OF MEASURES FOR CONTINUANCE OF THE SPECIES.

THE ENDOCRINE ORGANS AND REPRODUCTION

HISTORICAL NOTE

The Endocrine Organs. The subject of endocrinology belongs entirely to the present century, although, among the speculations of earlier writers was that of RUYSCH, about 1690, that organs such as the thyroid produced important substances which entered the blood stream. Also, in 1849 BERTHOLD induced comb growth in a capon by testicular grafting. The name "*internal secretion*" was, moreover, introduced by CLAUDE BERNARD in another connection, *i.e.* the glycogenic function of the liver: and in 1889 BROWN-SÉQUARD tried, upon himself, injections of testicular extract. The name *hormone* was introduced by BAYLISS and STARLING about 1905, and secretin was the first one to be recognised by a name.

The Thyroid was variously regarded during the 19th century as having to do with voice, as a regulator of blood flow to the head, or as of merely cosmetic purpose. It was first removed from animals about 1840 by ASTLEY COOPER, and SCHIFF, in 1856, showed that its removal (and with it the then undiscovered parathyroids) was fatal in most species. The relation of the thyroid to myxœdema was suggested by ORD in 1878; and REVERDIN in 1882, and KOCHER in 1883, showed that the condition of "*cachexia strumipriva*" followed thyroidectomy in man: HORSLEY, 1884, confirmed this on monkeys.

In 1884, SCHIFF resumed the subject, showing that the effects of thyroidectomy could be alleviated by a graft of thyroid, and MURRAY (1891) and FOX (1892) started thyroid therapy by oral administration in man. The success of this replacement therapy discredited the theory that the thyroid function was concerned with the removal of toxic substances produced in the body.

The discovery of iodine as an essential constituent of the thyroid, by GLEY in 1889, led, in 1895, to the finding by BAUMANN and ROOS, that the iodine was in organic combination. The active principle, thyroxine, was isolated by KENDALL in 1914, its structure was found by HARINGTON in 1926, and it was prepared synthetically by HARINGTON and BARGER in 1927.

The Parathyroids were discovered by SANDSTRÖM in 1880: in 1897, GLEY showed that many of the effects of "thyroidectomy" were attributable to the simultaneous removal of the parathyroids. These effects were at first thought to be due to accumulation of toxic substances (such as guanidine) which it was the function of the parathyroids to eliminate. A clear clue was given as to their functions when HOWLAND and MARRIOTT showed, in 1918, that the state of parathyroid tetany was accompanied by a fall in blood calcium. A hormone which raised blood calcium was extracted by COLLIP in 1925.

The Suprarenals were discovered by EUSTACHIUS in 1543, but no hint of a function came until ADDISON (1849) described the symptoms of the disease which bears his name. BROWN-SÉQUARD, in 1856, showed that their extirpation was soon fatal, and OLIVER and SCHÄFER, in 1894, found that extracts showed powerful physiological properties. This discovery focussed attention on the medulla, and ABEL and CRAWFORD discovered adrenaline in 1897. The affinity between the actions of adrenaline and those of the sympathetic system was demonstrated by LANGLEY in 1901.

Modern ideas regarding the function of the cortex began with the preparation, by SWINGLE and PFIFFNER in 1929, of an adrenal cortical extract capable of suspending the effects of adrenalectomy: the researches of KENDALL and others showed that a whole series of steroids were the responsible active ingredients.

The Pituitary was shown to be concerned in the disease acromegaly by MARIE in 1886, but the manner of its action was not established until about 1935, when the growth-promoting hormone was extracted by H. M. EVANS and others. Outstanding among other discoveries concerned with the anterior lobe were also those of H. M. EVANS and LONG (1921) and P. E. SMITH (1926), which demonstrated the important effects of specific anterior lobe hormones on the reproductive system, and on other endocrine organs.

The discovery, by OLIVER and SCHAFER, in 1895, of a pressor principle in the pituitary—

shown by HOWELL in 1898 to be confined to the posterior lobe, attracted attention fo many years to the posterior lobe, to the exclusion of the anterior: pressor and oxytoci principles of the posterior lobe were separated by KAMM and others in 1928. Hypophysectomy was first performed, on the cat, by MARINESCU, in 1892.

The Gonads. The striking consequences of removal of the gonads on sexual activity, and on the secondary sexual characters, has been known from the earliest times. Th explanation of the effects which was most popular in the 19th century was that o a "trophic" nerve effect. A chemical causation, due to the presence of an "internal secretion" was suggested, but BROWN-SÉQUARD'S experiments on administration o testicular extracts as a treatment of senile impotence which lent some colour to it, wer soon disregarded.

Success first attended experiments on the ovary, *e.g.* the effects of ovarian extracts on the œstrous cycle, and on ovariectomised animals: positive effects of such extracts wer obtained, for instance, by MARSHALL and JOLLY in 1906. The introduction of the vaginal smear technique by STOCKARD and PAPANICOLAOU in 1917 made easy the following of the œstrous cycle, and ALLEN and DOISY, in 1923, announced the presence in *liquor folliculi* of hormone(s) capable of inducing œstrus in ovariectomised animals. The first crystalline preparation of such a hormone was obtained by DOISY in 1929. The isolation and chemical identification of other hormones from the ovary and testis soon followed.

The male hormone androsterone was isolated from male urine by BUTENANDT in 1931.

Reproduction. Until almost within living memory, opinions on reproduction were largely of a mystical character, and rested on little experimental foundation.

In the Hippocratic writings (*ca.* 400 B.C.) the view was that seeds from all parts of the body of man and woman, flow together, and, uniting, form the fruit, which then develops.

ARISTOTLE (*ca.* 350 B.C.) believed that a male factor provided movement, and a female one material, the sex of the offspring depending on which predominated. GALEN (*ca.* A.D. 200) postulated the mixing of a male and female semen: heart and liver were formed from the blood of the mother, the brain from the male semen. FABRICIUS (1621) believed in the fecundating efficacy of an emanation from the semen, because, in birds, the egg was fertilised far from the site at which the semen was deposited. This theory of the *Aura seminalis* was further developed by HARVEY (1651): the uterus was brought by copulation to a state of ripeness, and acquired properties akin to those of the brain, and, like it, could "conceive": the stimulus to conception was the *Aura seminalis.* The conception was, as Hensen said, essentially a "ghostly" one. The notion of the *Aura* has lasted in the popular mind right up to comparatively recent times. The human egg, according to Harvey, was the result of the union of male and female semen, and was intermediate between living and non-living things.

For long a theory of "evolution" held sway. According to this, the egg contained the fully formed embryo, and, therefore, all the embryos of all the later generations. HALLER defended this view, which, however, received its death-blow at the hands of BLUMENBACH in 1784, and was replaced by the theory of "epigenesis."

From time to time there was fluctuation of opinion as to the contribution, by male and female, to the formation of the embryo. According to the one school the male element only provided a stimulus: according to the other, the male element was a seed, for the development of which the mother merely provided the soil.

Early notions on reproduction were also bound up with a belief in spontaneous generation. This latter belief was attacked by HARVEY (1651), and REDI, MALPIGHI, SWAMMERDAM and others in the 17th century brought additional evidence against it. SPALLANZANI'S researches in 1765 showed that many supposed instances of spontaneous generation of lif could be explained by the presence of viable, though dried, spores or seeds, but it was not until PASTEUR'S work in the latter half of the 19th century that the idea of spontaneous generation was almost universally abandoned by educated people. Understanding of the nature of reproduction evidently awaited the application of adequate histological technique.

The ovary, called *testis muliebris* by GALEN, was renamed *ovaria* by STENSON in 1664, and in 1672 DE GRAAF saw its follicles, and regarded them as eggs. The actual ovum was seen in 1827 by V. BAER, who also traced its course, in the bitch, along the Fallopian tube into the uterus, though De Graaf had also already seen them in the Fallopian tubes and uterus.

Spermatozoa were discovered by a Dutch student HAM and were studied in detail by LEUWENHOECK in 1678; the name was given because (even till 1835) they were thought by some to be small parasitic animals; the error of this was exposed by KÖLLIKER in 1841 when he followed their development in the testis.

The first demonstrations of the feasibility of artificial insemination (JACOBI, 1764, SPALLANZANI, 1786) were supposed to have shaken the belief in the *Aura seminis.* The essential feature of fertilisation as the entry of the spermatozoon into the ovum was appreciated by BARRY in 1853, and was of the greatest biological importance.

The relation of ovulation to menstruation was first recognised by BISCHOFF in 1844. The physiology of erection was founded on the work of J. MULLER (1838), ECKHARD (1863–76) and others, and is of classical significance, chiefly in relation to the vaso-dilator action involved.

The considerable body of present-day knowledge concerning the physiology of reproduction is largely the product of the last few decades of investigation.

A. THE ENDOCRINE ORGANS

CHAPTER L

WHEN, in a primitive organism, the reaction of distant parts of the body to a change occurring in any one part depends on the diffusion of some substance from the stimulated part, the total reaction must require a considerable time for its full development. A much more effective method of correlation was acquired by the evolution of a nervous system, but this second mode of correlation did not do away with the necessity for the more primitive method. Even in the higher animals, when sustained, rather than evanescent reaction is required, we often find response to some change in distant parts of the body, the message being chemical and not nervous in character (*e.g.* the secretin mechanism for pancreatic secretion).

When we speak of the chemical correlation of the activities of the different parts of the body, it is important not to confuse processes which have little or nothing in common. In one sense, we may say that every cell in the body is chemically connected with, and dependent on, all the other cells in the body. Thus, one organ may receive and modify any substance or foodstuff so as to prepare it for more ready assimilation by other tissues. It may shield other tissues from the poisonous effects of certain waste products, either by converting them into harmless substances or by excreting them from the body.

We are probably justified in treating apart a group of phenomena in which the substance transmitted from one part of the organism to another is significant almost entirely as an excitatory or inhibitory agent, and has little or no value as a source of energy. In previous chapters, we have already come across several examples of correlated activities of different tissues, effected by chemical means. Thus *secretin*, which is heat stable and diffusible but is easily destroyed by oxidation, is absorbed into the blood and carried to the pancreas, where it acts as a specific stimulus for the secretory cells. This was the first discovered example of chemical co-ordination, *i.e.* a response effected by chemical means rather than by impulses passing along nerves. When the adaptation to a change at A consists in increase or diminution in the activity of an organ B, the change in B can be evoked by the production at A, as a direct consequence of a stimulus, of a specific chemical substance, which passes in the circulating blood to B, where in its turn it will produce the required state of action. Such chemical messengers may be called in general, *autacoids ;* when the effect is to cause excitation they may be called *hormones*, from ὁρμάω, "I excite." * The same autacoid may exert a stimulating action on one organ and a depressant one on another, however. Usually they are now all called hormones, irrespective of the direction of their action. It may be interesting here to consider what must be the characters of substances which fulfil the part of chemical messengers.

Since they must be carried by the blood stream to the reacting organ, they must be susceptible of passage through the walls of the capillaries, and rapidly so, if they are to excite a reaction within a short time. Such substances would need to be of small molecular size, and actually, some of them have a comparatively simple chemical and molecular constitution. Some, however, are of high molecular weight, and act slowly.

Chemically the hormones fall into three groups, (1) *Phenolic derivatives*, *e.g.* adrenaline and thyroxine, (2) *Steroids*, *e.g.* the sex hormones and the adrenal cortical hormones, and (3) *Proteins*, *e.g.* the hormones of the pituitary, pancreas and parathyroid.

---

* STARLING. Croonian Lect. R.C.P., 1905 ; JAYLE. *Bull. de l'académ. de Méd.*, 1937, 117, 73.

Hormones fall into two rough groups according to whether their actions are fugitive or lasting. In those of the former group, the chemical messenger must excite a state of activity (or of inhibition) in response to a change in some other part of the body, *e.g.* when adrenaline is thrown out from the suprarenals in states of sympathetic excitation. When the primary change passes away, the action of the hormone should also disappear. On this account it is necessary that the hormone should either be susceptible of easy destruction, by oxidation or otherwise, in the fluids of the body, or be readily excreted.

In the other group of hormones, which are more stable, the effect is protracted, appearing and disappearing slowly, as, *e.g.* is the case with the hormone of the thyroid, with the sex hormones, and with some of the pituitary hormones.

In the form in which they are liberated in the body, the hormones must not be *antigens, i.e.* their presence in the blood stream must not involve the production of an antibody. If this were the case, the hormone, on entering the blood stream, would meet its antibody and would be unable to exert any effect on the appropriate reacting organ.

The production under experimental conditions of anti-hormones to the active principles of the pituitary or parathyroid, when these are injected, is an apparent exception. The hormones in question are proteins, and the production of the anti-hormones is only observed, or is most evident, when the pituitary material is derived from a different species of animal from that receiving the injection.

The hormones show a fairly well-defined *specificity* as regards the tissues on which they act, but not usually as regards the species of animal. Thus adrenaline from any species of animal has the same actions in general on the tissues of any other species. But the tissue specificity is not absolute : thus sex hormones of one sex may produce some effects on the other sex.

There is one class of organs in which the production of hormones is the only function with which we are acquainted. These are included under the term *endocrine organs*. As examples, we may cite the suprarenal bodies, the thyroid and parathyroids and the pituitary body. Structural proximity, however, does not mean relation of function. Thus we have the cortex and medulla of the suprarenal, each with distinct endocrine functions ; similarly the anterior and posterior parts of the pituitary body, while the thyroids and parathyroids are often as close in situation as they are remote from one another in function. In other cases, an organ may have multiple functions. Thus, the pancreas gives an external secretion, the pancreatic juice, and an internal secretion, insulin, which, passing into the blood, exercises an important influence on the metabolism of carbohydrate. Other instances of these chemical correlations may be cited, *e.g.* the gastric hormone. In the reproductive system, we have many instances of such chemical correlations.

Methods of Investigation. The investigation of the functions of the endocrine organs consists essentially in studying the effects, on the various functions of the body, of augmenting and of diminishing the amount of active principle of the organ in question. There are several directions along which this information may be obtained. Thus, in certain diseases, the endocrine organ concerned produces less of the active principle than it does in the normal state, or, conversely, in conditions associated with overgrowth or overaction of the gland, it may produce more. Careful study of the clinical effects of the hypo- or hyper-function of the particular endocrine organ therefore throws light on the functions which it normally performs.

The effects of lack of the endocrine principle can be studied, in experiment on animals, by the removal by operation of the gland in question, such operations being sometimes performed in stages, so that the resultant change is a gradual one. Then, having produced the effects of lack of the active substance, we may attempt to make good this lack by the administration of the substance, either by feeding with the gland, by administering extracts of it intravenously or by subcutaneous injection (insoluble substances being implanted subcutaneously, or injected suspended in oil); or by the surgical grafting of a portion of the gland from another animal of the same species. We can also, of course, study the effects of similar administration of the active principle to the normal animal or man, and thus learn the physiological effects of an excess of the hormone.

Next, chemical investigation seeks to isolate and to ascertain the nature of these active principles, with the view to their ultimate synthesis in the laboratory, an end which has already been achieved in several instances. It is also essential to find some method by which the amount of the active principle present in a tissue or extract may be estimated. This is usually a method of biological assay, as, *e.g.* in the case of the standardisation of insulin preparations, as chemical methods are not delicate enough. A unit of biological activity is decided upon (often by international agreement), and, when possible, this is related to a known weight of the pure active substance.

The presence, in extracts of tissues, of substances which exert definite physiological effects when administered in appropriate ways, is, however, no proof that the organ in question normally produces that substance for the benefit of the whole body, or that the substance is ever passed, as such, from the organ into the blood or lymph. Strict proof would require evidence that the organ in question, (*a*) contained an active chemical substance, which (*b*) acts upon some other parts of the body, so that (*c*) definite effects would ensue, if the tissue which forms this active principle were all removed, but (*d*) these effects would disappear if the active principle was suitably administered; (*e*) the active substance should be demonstrable at certain times in the blood or in the lymph leaving the organ in which it is produced.

Another important line of enquiry is that which concerns the paths by which the hormones are produced in the body—their precursors, mode and place of synthesis, etc., and also the means by which they are removed or destroyed in the body.

Finally, it is important to study the means by which the formation or liberation of the hormones is effected. In general the liberation of hormones into the blood stream appears to be a continuous process, and to be independent of influences of the nervous system. But there are important exceptions to this, *e.g.* the release of adrenaline into the blood. Usually the regulation of hormone secretion is controlled by chemical factors, such as the presence of other hormones or blood constituents.

THE SUPRARENAL BODIES

The suprarenal bodies, or adrenals, in mammals are two small masses lying on or near the upper or oral pole of the kidneys. They consist of two parts, the cortex and the medulla. The cortex is composed of cells arranged in columns or in a reticular fashion; in the female in some species, and notably in mice, the cortex is much larger than in the male. It is doubtful whether any nerves supply the cortex. According to the arrangement of its cells the cortex is divided into three zones, the zona glomerulosa, zona fasciculata, and zona reticularis. The cells themselves are distinguished by the large amount of granules they contain, which probably consist of lecithin and sterol compounds. These granules stain with osmic acid, and are dissolved out by turpentine. The cortex cells stain black with silver nitrate on account of their high content of vitamin C. In some animals, *e.g.* the guinea-pig, the cells, especially towards the inner part of the cortex, contain many yellow pigment granules.* In others, especially in young female and castrated male mice, the inner layers of the zona reticularis are increased in thickness, and consist of small deep-staining cells: this is called *the X-zone.* It shows degeneration in older, and in pregnant, females.

The medulla, much less extensive than the cortex, presents irregularly shaped cells, the outlines of which are but slightly marked. These cells contain granules which stain darkly with chromates and give a green colour with salts of iron. It is hence easy to delimit the area of the medulla in any section of a gland which has been hardened in a fluid containing chromates. The substance giving this reaction is known as chromaffine substance.[*] Fine osmic-staining granules are also present in abundance in these cells, but are *not* dissolved out by turpentine. All these reactions are due to the presence of adrenaline.

---

* Cramer. *J. Physiol.*, 1918, **52**, Proc. viii.

The suprarenals are richly supplied with blood, especially in the medullary part, the cells of which impinge directly on the endothelial lining of dilated capillaries. The blood supply is probably the most abundant in the whole body. They also receive an abundant nerve supply from the sympathetic system (splanchnics), and some from the vagus, and perhaps from the phrenics.[†] The nerves form a thick meshwork in the medulla; this contains also a few ganglion cells, which may either be isolated, or combined to form small ganglia.

A study of the development of the suprarenal glands shows that we have here to do with two distinct tissues, which have different origins and different functions in the body. Whereas the cortex is derived from that portion of the mesoblast, the "intermediate cell mass," from which the mesonephros is also developed, the medulla is produced by an outgrowth from the sympathetic system, and may indeed be said to consist of profoundly modified nerve cells. In many fishes, these two elements of the suprarenal gland remain separated throughout life, the cortex being represented by a series of paired inter-renal bodies lying on the front of the spinal column, and the medulla by a number of collections of chromaffine cells, lying in close juxtaposition to the spinal nerves.

In some animals, accessory suprarenals, in which both cortex and medulla may be represented, are not infrequent. In all animals we find masses of tissue, the so-called para-ganglia, in close association with the sympathetic system, which present the chromaffine reaction typical of the medulla. Since a watery extract or decoction of these bodies has the same influence, on injection into the blood stream, as an infusion of the medulla of the suprarenal body itself, we are probably justified in regarding these bodies as equivalent to accessory medullary portions of the suprarenal. They have the same origin, the same staining reactions, and the same physiological effect as the latter.

Scattered portions of cortical tissue are also found in many animals, usually near to the gonads.

FUNCTION OF THE CORTEX

Extirpation of one suprarenal (single adrenalectomy) has no effect beyond that of sometimes leading to hypertrophy, especially of the cortex, of the other one. This may be due to a mutual reaction between suprarenal cortex and anterior pituitary, resulting in increased formation by the latter of the adreno-corticotropic hormone. The extirpation of both suprarenals is fatal in a few days if all accessory cortical tissue is also removed. In the cases where animals survive double extirpation, small masses of accessory cortical substance have been found embedded in the kidney or elsewhere in the neighbourhood of the suprarenals or gonads. Old male rats, on suitable diets, and some rabbits, may survive double adrenalectomy indefinitely because of accessory tissue left behind. To cats, dogs, young rats and guinea-pigs, however, the operation is fatal in, at most, a few days; death is accelerated in cold environment. It is certain that the rapidly fatal effect of extirpation of both suprarenals is to be ascribed to the removal of the cortex and not of the medulla. After recovery from the operation the basal metabolism of the animal, as measured by oxygen consumption, remains normal for a few days, and then, when symptoms of weakness appear, falls considerably.[‡] When death occurs it is sudden, and is preceded by rapid rise of blood urea and non-protein nitrogen, loss of appetite, imperfect fat and carbohydrate absorption, weakness of muscles, lowering of arterial pressure and of blood sugar, fall of body temperature, collapse, and a shock-like condition, with great concentration of the blood. It was shown by Loeb [§] that there is an increased rate of loss of sodium salts by the kidneys after adrenalectomy; the general dehydration and reduction of blood volume is to some extent a consequence of this lowered blood sodium. Since the urea clearance is reduced, and the excretion of other substances also disturbed, it would appear that there may be some primary interference with kidney function.[||] Thus blood potassium, creatinine, phosphate and sulphate are raised, while excretion of injected phenolsulphonephthalein is reduced. Any factor which augments loss of water from the blood aggravates the condition of circulatory collapse, *e.g.* it can be evoked very easily in an adrenalectomised animal by the administration of a small dose of histamine, to which such animals are abnormally sensitive. A similar loss of fluid and electrolytes from the blood occurs in wound shock.* Conversely, the injection of salt solution into the blood stream, or administration of NaCl by the mouth improves the condition for a while.

---

[*] Vincent. *Proc. Roy. Soc.*, 1910, B. **82,** 502.

[†] Teitelbaum. *Quart. Rev. Biol.*, 1942, **17,** 135.

[‡] Webster, Pfiffner and Swingle. *Amer. J. Physiol.*, 1932, **99,** 710.

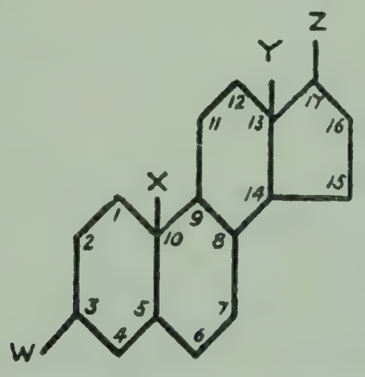
[§] Loeb *et al.* *J. Exp. Med.*, 1933, **57,** 775.

[||] Marshall and Davis. *J. Pharmacol.*, 1916, **8,** 525.

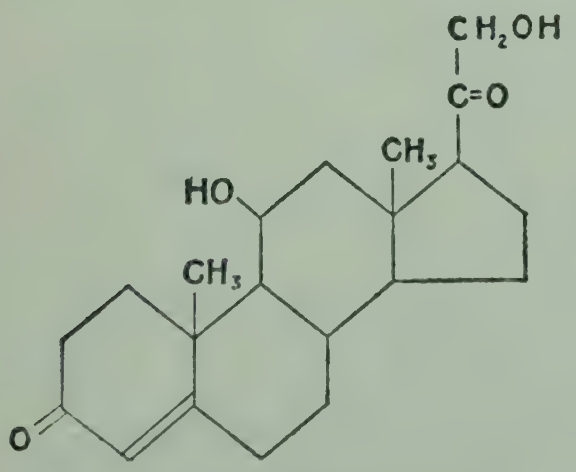
Addison in 1855 drew attention to the coincidence of degenerative destruction of the suprarenals with a disease which since that time has been known as Addison's disease. The three cardinal symptoms of this disorder are (1) bronzing of the skin and mucous membranes, (2) vomiting, (3) excessive muscular weakness and prostration. The disease is almost invariably fatal if untreated. Addison's observations have been amply confirmed since that time, but we are not yet in a position to account for the occurrence of all these symptoms as a result of interference with the suprarenals since it has been impossible experimentally to reproduce the bronzing which is so characteristic of Addison's disease.

Among many attempts to isolate from the cortex a substance capable of prolonging indefinitely the life of adrenalectomised animals, the first clear success was reached by Swingle and Pfiffner in 1929. They extracted fresh ox suprarenals with alcohol, concentrated the extract and re-extracted it with benzene and various other solvents, and finally obtained a water-soluble extract of sufficient purity to permit of administration by intraperitoneal injection.† Daily injections of small doses of extract suffice to maintain in health a previously adrenalectomised dog. Large doses given by mouth are effective.‡ When the administrations are withheld, the animal rapidly goes downhill, and dies in a few days: but even moribund animals are very soon restored if the dosage is recommenced, and sodium chloride given.

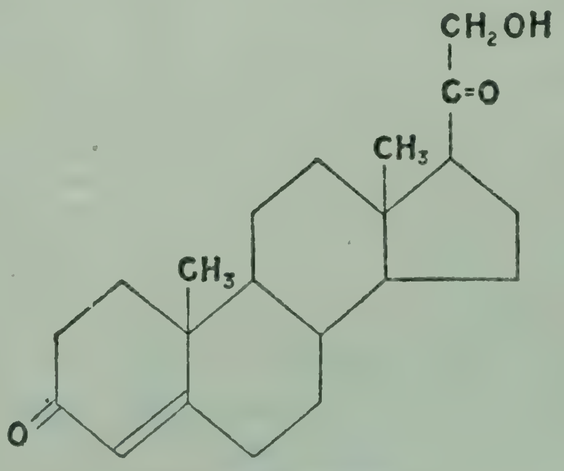
Cortical Hormones. From the researches of Kendall,§ and others, it has been found that some twenty sterols are present in the suprarenal cortex. The cholane ring which represents the steroid skeleton is :—



Chief among the cortical steroids are *Corticosterone*, *Desoxycorticosterone*, *Dehydrocorticosterone*, and, in small amounts, *Progesterone* and *Androsterone*



Corticosterone.



Desoxycorticosterone.

---

* MacKay *et al.* *Amer. J. Physiol.*, 1935, **112**, 581.

† Swingle and Pfiffner. *Amer. J. Physiol.*, 1931, **96**, 153, 164, 180 ; **98**, 144.

‡ Britton, Flippin and Silvette. *Am. J. Physiol.*, 1931, **99**, 44.

§ Kendall *et al.* *J. Biol. Chem.*, 1937, **120**, 719.

(the last two also being known among the sex hormones). They differ in their activity in replacing the hormone lost by adrenalectomy, though at least six of them are more or less active. They are almost certainly manufactured, in the cortical cells, from cholesterol.

The cortical steroids fall into two groups, one typified by corticosterone having a keto- or hydroxy-oxygen at C-11, and one typified by progesterone and desoxycorticosterone, which lack such a group. Their actions are different.

The essential activity of the cortex is probably to be attributed to the three hormones, Corticosterone, Desoxycorticosterone and Dehydrocorticosterone. In some ways the most active, though not most abundant, of the cortical steroids is desoxycorticosterone, which is now made synthetically and placed on the market as its acetate. The desoxycorticosterone type is more potent than corticosterone in affecting the salt and water balance, though it has less effect on carbohydrate and protein metabolism. The effect on carbohydrate metabolism is possibly due to its relation to phosphorylation.

A steroid compound, or complex, containing ascorbic acid, and behaving like dehydrocorticosterone, has also been isolated.*

Clinically, the solid hormone may be implanted under the skin in cases of Addison's disease.†

Desoxycorticosterone can restore an adrenalectomised animal or a patient with Addison's disease to a reasonably normal state of health. Testosterone can also cause improvement in strength, positive N-balance, and lowering of serum K‡; so can other cortical sterols, though in some cases, *e.g.* with androsterone or progesterone, the amounts needed are relatively large.

In muscular exercise, and as part of a generalised response of the body to damaging influences, such as trauma, cold, reduced atmospheric pressure, and certain toxic substances, there is evidence of increased activity of the suprarenal cortex, and the administration of cortical hormones in large doses has been used in combating surgical shock.

The cortical hormone, which is probably present in the lipide granules of the cells, is released into the blood stream, and can be assayed in the venous blood leaving the gland.§ The assay is made by determining the prolongation of life of suprarenalectomised rats kept in cold surroundings when the plasma of the sample of blood is injected: the results show that a large amount of the hormone is continually being released, the amount from one suprarenal in the dog being estimated as equal in twenty-four hours to the commercially extractable hormone content of no less than 17,300 g. of gland. It is evident, therefore, that either the usual methods of extraction are very inefficient, or else that new hormone is continually and rapidly being produced in, and liberated from, the cortex. The rate of liberation is independent of the rate of blood flow through the gland, or of the arterial blood pressure, but is augmented to several times its resting value by the injection of physiological amounts of adrenaline,‖ and hence also by splanchnic excitation. It is also accelerated by addition to the blood of K· ions, A.T.P. or creatine phosphate.¶ There is no direct action of the sympathetic on the cortex, but the indirect effect, through production of adrenaline, provides a means of supplying, with that adrenaline, a second hormone which plays an important part in enabling the body to withstand stress. The hormone rapidly disappears from the blood and so is not detectable in any blood but that of the suprarenal vein. This is so even if the alimentary canal, liver and kidneys have been removed: it can therefore in all probability be removed by the tissues generally, but normally part is excreted, in active or inactive form, in the urine.

---

* Lowenstein and Zwemer. *Endocrin.*, 1946, **39**, 63.
† Dryerre. *B.M.J.*, 1939 (i.), 971.
‡ Talbot, *et al.* *J. Clin. Inv.*, 1943, **22**, 583.
§ Vogt. *J. Physiol.*, 1943, **102**, 341.
‖ Vogt. *J. Physiol.*, 1944, **103**, 317.
¶ Vogt. *J. Physiol.*, 1949. *Proc.* 45P.

Relations between Gonads and Adrenals. In view of the fact that some of the adrenal corticosteroids exhibit androgenic or œstrogenic properties, or are even identical with sex hormones, it is not surprising that close and complex relations should exist between the cortex and the sex organs and secondary sex characters. These relations can here only be touched upon, but perhaps the most striking illustration is that of the disturbance of secondary sex characters seen in tumours or hyperactivity of the adrenal cortex. According as the output of the cortex in these cases is preponderatingly androgenic or œstrogenic, we may get precociously sexed boys, masculinised girls, women with virilism, or feminised men. The most frequent cases are of virilism in women.

Hypertrophy, or a tumour, of the suprarenal cortex in children is associated with premature sexual maturity * and in women with masculinisation and hirsutism (and also often with a basophile pituitary tumour). Many cases of masculinity (virilism) in women have been dramatically cured by removal of an adrenal tumour ; the male hair may then be removable a few days after operation.

In such cases the inner layers of the cortex were found by Vines to contain granules stainable by ponceau-fuchsin. This material is absent from normal suprarenals, except in fœtal life, up to the end of the first half. The fuchsinophile material appears sooner and persists longer in male than in female fœtuses, and is believed to indicate the presence of androgenic sex hormones. The occurrence of virilism in women seems to be due to the obtrusion of male qualities which in the normal endocrine balance of the female, are held in abeyance. The corticosteroids essential to life have little effect on the sex functions when given to intact animals ; desoxycorticosterone, however, has been shown to have a progesterone-like action, and progesterone itself is present in the cortex. Substances having actions similar to those of cortical sterols have been recovered from the urine of mares and cows.† By utilising a colour reaction with *m*-dinitrobenzene and KOH, a comprehensive estimation may be made of all the 17-ketosteroids in the urine ‡ : these include several adrenal cortical and androgenic steroids. Normal men excrete 10–15 mg. and women 6–10 mg. daily : in children prior to puberty the amounts are much smaller ; in adrenal cortical tumours the amounts of 17-ketosteroids may rise to 200 mg. daily.§

Relations of Cortex with Pituitary Anterior Lobe. After hypophysectomy the inner layers of the cortex, first the zona reticularis and then the zona fasciculata shrink and degenerate, and these changes can be prevented by the injection of suitable extracts of the anterior lobe—the so-called adrenocorticotropic hormone. Hypertrophy of the inner cortical layers can be induced by the injection of these extracts into normal animals. The influence of the cortex on the metabolism of carbohydrate, fat and protein is the most affected by the adrenocorticotropic factor, from which it may be suspected that it is the inner layers of the cortex which secrete the hormones that are concerned with the control of these aspects of metabolism. The outer, and more newly-formed cells of the zona glomerulosa may either secrete enough hormone, or a hormone of the right kind, to exercise control over the salt and water regulation of the body after hypophysectomy.

Nature of Action. Various explanations have been advanced to attempt to explain why the suprarenal cortex is essential to life, and to find in what the fundamental action of the cortical hormones consists. Apart from any theories, it is evident that they must have some action which affects processes that are quite indispensable to life ; no doubt some of the effects are secondary to this action. Opinions differ as to what the fundamental action is, and depend very much on the clarity with which results emerge.

---

* Bulloch and Sequeira. *Trans. Path. Soc., London,* 1905, **66,** 189.

† Selye and Dosne. *Lancet,* 1940, **239,** 70.

‡ Callow *et al.* *Bioch. J.,* 1938, **32,** 1312.

§ Callow and Crooke. *Lancet,* 1944, **246,** 464.

The striking excretion of sodium from the body after adrenalectomy, together with the accumulation of potassium and loss of water, have led to the theory that the fundamental action of the hormones is related to the control of the salt and water balance, and that the other effects, such as those on carbohydrate metabolism, are in some way related to this. But these effects on salts and water are chiefly due to the desoxy-hormone.

The tissue chlorides certainly play some important part in general metabolism, as may be shown by the experiment of acute chloride depletion. A dog is given a large intraperitoneal injection of 5 per cent. glucose, say 1 l. ; much of the blood chloride diffuses into this, and three hours later the fluid is withdrawn. It is then found that the metabolic rate (oxygen usage) of the animal is low, but is restored to normal by intravenous NaCl.[*] The NaCl is said to play a part in facilitating transfer of oxygen from hæmoglobin to the tissues.

Another view is that the fundamental effect is on carbohydrate metabolism, and as we have seen in Chapter XLVI, this effect is certainly quite pronounced. This effect is due to the corticosterone type of hormone. It consists in a depression of glucose usage by the tissues together with accelerated glucose formation from protein or amino-acids, and increased glycogen storage. A way in which the protein metabolism may be affected may be by an effect on liver arginase, which is increased by corticosterone.

Units. Various units of activity of the cortical principles have been proposed. Thus, the Pfiffner " dog unit " is the minimal daily dose per kg. body weight which will keep an adrenalectomised dog constant in body weight and with a constant low blood urea. If 0·3 g. kg./day of NaCl is given with the diet, the unit is equal to 0·5–1·0 g. of fresh gland, or 0·015–0·020 mg. desoxycorticosterone kg./day. If sodium salts are not administered, at least ten times as much will be needed.

The suprarenal cortex also has a high content of ascorbic acid, which is probably due to its storage there, perhaps combined with a steroid.

FUNCTIONS OF THE MEDULLA

The medulla can be destroyed by operation, or by insertion of a small tube of radium, without causing death. It seems that the functions of the medulla, if essential to life, can be more or less effectively maintained by the other chromaffine tissue found at the back of the abdomen. But there is no conclusive evidence that the chromaffine material is, in fact, essential.

Schäfer and Oliver in 1894[†] found that an aqueous extract of the suprarenal bodies, when injected into the circulation, caused a great rise of blood pressure, brought about chiefly by generalised vaso-constriction. The substance responsible for this was limited to the medulla, and later on Abel [‡] isolated an active substance to which the name of *adrenaline* [§] was given. Long afterwards *nor*-adrenaline was shown to be present too (formulæ p. 361).

Adrenaline gives a blue colour with phosphotungstic acid,[‖] or with arsenomolybdic acid ;[¶] the latter can be employed for its quantitative estimation, but it is more often assayed in biological fluids by a biological method, *e.g.* by the relaxation of a loop of isolated intestine. An even more delicate test is the relaxation of the non-pregnant rat uterus suspended in Tyrode solution containing excess of KCl (0·2 p.c.) : this can detect a concentration of 1 in 10^9 [**]

[*] Cohn and Soskin. *Amer. J. Physiol.*, 1943, **139**, 80.

[†] Schäfer and Oliver. *J. Physiol.*, 1895, **18**, 230.

[‡] Abel. *Amer. J. Physiol.*, 1900, **3**.

[§] Also called adrenine or epinephrine.

[‖] Baker and Marrian. *Biochem. Journ.*, 1927, **21**, 1005.

[¶] Shaw. *Bioch. J.*, 1938, **32**, 19.

[**] King. *J. Physiol.*, 1949, **109**, *Proc.*, in press.

It may be distinguished from nearly related compounds by the green fluorescence, shown in ultraviolet light, when alkalised.*

The natural form is lævorotatory and is physiologically about twelve times as active as is the dextro-variety.† Both dextro- and lævo-forms have been prepared synthetically. The general chemical and physiological properties of *nor*-adrenaline resemble those of adrenaline, the physiological differences being rather quantitative than qualitative.‡

It is accepted that adrenaline and *nor*-adrenaline as prepared from the suprarenal are identical with the active hormones which these organs can release into the blood stream, and with the synthetic compounds. When we speak of a release of " adrenaline " from the suprarenals, we mean a release of both substances, but the conditions under which one or the other might preponderate are as yet uncertain.

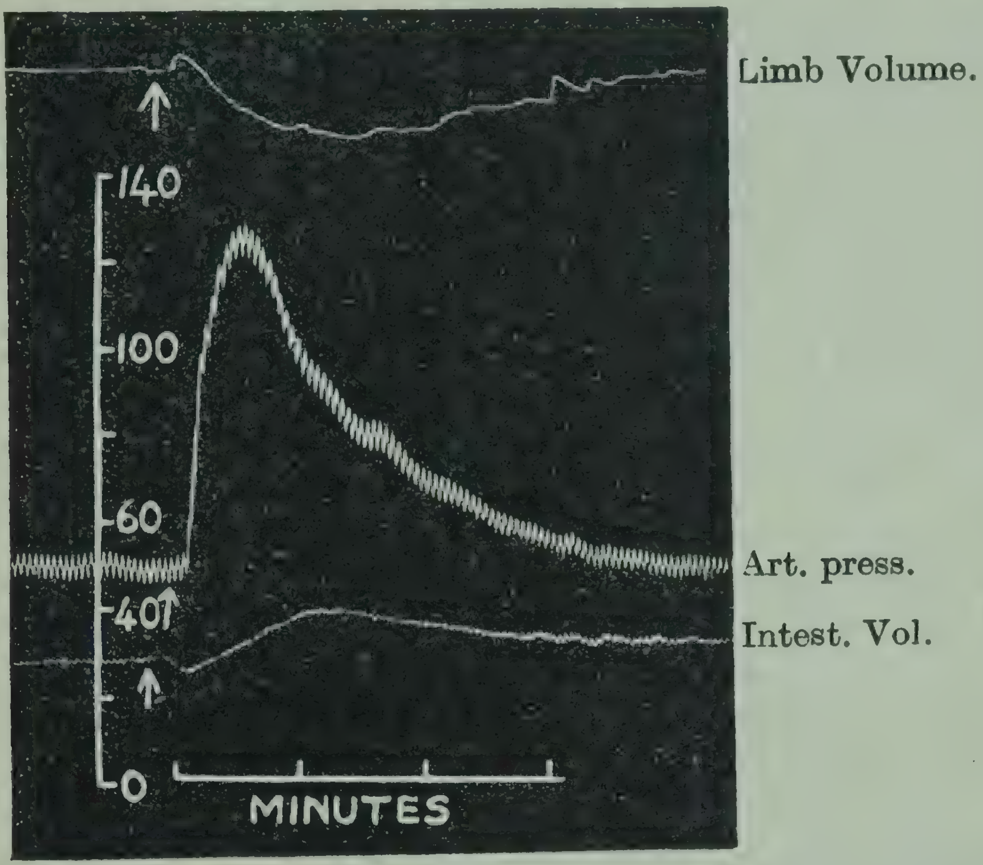


Fig. 639. Effect of Intravenous Injection of Adrenaline on Limb Volume, Arterial Pressure and Intestine Volume of Decerebrate Cat. At arrow 0·01 mg. adrenaline; passive dilatation of intestine. (G. A. Clark. *J. Physiol.*, 1930, 69, 181.)

Adrenaline, though almost inert when given by mouth, is active in minute amounts on injection into the general circulation, injection of 0·0025 mg. per kg. body weight sufficing to evoke an effect upon blood pressure. On injecting it into the circulation there is immediately a rise of blood pressure which, if the vagi are intact, is only moderate in amount, since it is accompanied by a slowing of the heart. This latter is due to an excitation of the vagus secondary to the rise of blood pressure, and is not due to direct action of the drug on the vagus centre except that it increases its reflex excitability.§ (The slowing is more evident with *nor*-adrenaline.) If the vagi be divided, the injection of adrenaline, even after destruction of the central nervous system, evokes a huge rise of pressure, which may amount to 300 mm. Hg. The most intensely constricted vessels are those of the skin; in the splanchnic area the effect is smaller, so that if the rise of pressure is large, these vessels may become passively dilated (Fig. 639).

Adrenaline injected into the circulation arouses those activities which can normally be excited by stimulation of the sympathetic system.‖ With *nor*-

* Gaddum and Schild. *J. Physiol.*, 1933, **80**, 9P.

† Cushny. *J. Physiol.*, 1908, **38**, 259.

‡ West. *J. Physiol.*, 1947, **106**, 418.

§ Stella. *J. Physiol.*, 1932, **77**, 68.

‖ Langley. *J. Physiol.*, 1901, **27**, 237.

adrenaline the parallel is even closer (*v.* p. 361). After adrenaline injection the liver glycogen is rapidly discharged, and hyperglycæmia and glycosuria result.* Potassium is also liberated from the liver.† The glycogen of muscle is also discharged, being converted into lactic acid.‡ Fatigue of skeletal muscle is postponed. The heat production and oxygen usage are increased.§ In the head, it causes dilatation of the pupil, secretion of saliva, and erection of the hairs. The dilatation of the pupil is much accentuated after previous removal of its sympathetic supply by excision of the superior cervical ganglion.‖ On the heart it has a strong augmentor and accelerator influence, greatly increasing its oxygen utilisation,¶ so that it beats more effectively, and empties itself more completely in systole, even against the enormously increased resistance offered by the constricted arterioles. The bronchi and bronchioles are fully relaxed ** ; respiratory movements are often temporarily inhibited (adrenaline apnœa : this is a secondary effect of the rise of arterial pressure on the carotid sinus). On the lung vessels, adrenaline has probably a slight constrictor influence. With regard to the vessels of the brain, we find the same slight constriction as in the case of excitation of the sympathetic nerves to this organ, In the abdomen, it causes complete relaxation of the musculature of the cats and the human uterus, of the stomach, and of the small and large intestines but contraction of the ileocolic sphincter, and of the rabbit's uterus. On the bladder, its effect varies according to the animal studied, but in every case is identical with that obtained by stimulating the hypogastric nerves. It has been shown that very small doses of adrenaline usually cause a definite *fall* of blood pressure in anæsthetised animals. This is probably due to relaxation of the arterioles,†† particularly in muscle,‡‡ outweighing the vasoconstrictor effect elsewhere ; *nor*-adrenaline only produces constrictor effects, hence its pressor action is more powerful than that of adrenaline. Its action on the pregnant cat's uterus is also more powerful, but otherwise its actions are generally weaker, than those of adrenaline.

The motor effects of adrenaline are usually reversed after previous injection of ergotamine, the active alkaloid of ergot, or of dibenamine. These drugs cause paralysis of all adrenergic mechanisms. Excitation of the splanchnic nerves, or injection of adrenaline, will now bring about a fall of blood pressure, due to dilatation of the vessels in the splanchnic area.§§ The substance 933 F. (piperidomethylbenzodioxane) is also an anti-sympathomimetic substance, which reverses the hypertension caused by adrenaline, though it does not paralyse adrenergic nerves.

The point of action of the adrenaline appears to be in the muscular or glandular tissues themselves, since its effects may not only be obtained after destruction of the sympathetic nerves, but even obtained (and in an exaggerated degree) after time has been allowed for the peripheral (post-ganglionic) fibres to degenerate as a result of extirpation of the corresponding ganglia. We have already dealt (p. 360) with the probability that the "adrenergic" nerves produce their effects by the liberation at their endings of some related substance, probably *nor*-adrenaline.

The splanchnic nerve is the secretory nerve to the suprarenal medulla. The fibres end without the interposition of any ganglion cells on their course,‖‖

---

* Blum. *Pflüger's Arch.*, 1902, **90,** 617 ; Hester and Wakeman. *Arch. f. exp. Path.*, 1902, **169,** 478.

† D'Silva. *J. Physiol.*, 1934, **82,** 393 ; 1949, **108,** 218.

‡ Cori and Cori. *J. Biol. Chem.*, 1928, **79,** 309.

§ Boothby and Sandiford. *Amer. J. Physiol.*, 1919, **51,** 200.

‖ Meltzer and Auer. *Amer. J. Physiol.*, 1904, **11,** 28.

¶ Lovatt Evans and Ogawa. *J. Physiol.*, 1914, **47,** 446.

** Thornton. *Q. J. Exp. Phys.*, 1932, **21,** 305.

†† Dale and Richards. *J. Physiol.*, 1918, **52,** 110.

‡‡ Clark. *J. Physiol.*, 1934, **80,** 429.

§§ Dale. *J. Physiol.*, 1913, **46,** 291.

‖‖ Elliott. *J. Physiol.*, 1913, **46,** 285.

the cells of the medulla themselves being modified sympathetic ganglion cells. Stimulation of the peripheral end of the splanchnic nerve causes a discharge of adrenaline and *nor*-adrenaline into the blood stream. This discharge accounts for the secondary rise, often accompanied by quickening of the heart, observed on a blood-pressure tracing as the result of stimulating the splanchnic nerve.* Through the splanchnic nerves, a discharge of adrenaline can be excited by many general conditions, such as anoxæmia,† pressure on the brain, puncture of the fourth ventricle,‡ administration of anæsthetics, or mental disturbances such as excitement or fright. Such a discharge is an important element in the adaptation to environmental stress, and enables the animal to react for the preservation of its life, either by offence or by flight. If one splanchnic nerve be cut before the administration of anæsthetics or the induction of a condition of irritation or fright, the suprarenal gland on the corresponding side will be found to contain two or three times as much adrenaline as the gland which has been left in connection with the central nervous system. It is stated that no such condition of exhaustion can be produced by electrical stimulation of the peripheral end of the divided splanchnic. This may be because the splanchnic fibres are not very suitably excited by faradic stimulation, and so the recuperative process in the medulla is able to keep pace with the depletion.

The splanchnic endings in the medulla are cholinergic, and small intra-arterial injections of acetyl choline into the suprarenal cause a large discharge of "adrenaline".§

Cannon and his co-workers‖ demonstrated the liberation of adrenaline from the suprarenals, by observing the acceleration of the denervated heart. Cats were used in these experiments, and the heart was denervated by removal of the stellate and other thoracic sympathetic ganglia and section of the vagi in the neck. Such animals, in states of asphyxia, fear, pain or rage showed pronounced cardiac acceleration, but if the splanchnics were divided, the effect was very slight or absent. Reflex secretion of adrenaline could also be shown in this preparation by stimulation of the central end of the sciatic nerve.

Similarly, Kellaway, using a cat with the pupil of one eye deprived of its sympathetic supply, showed that the hypersensitive denervated pupil dilates in even mild degrees of asphyxia or oxygen lack.

Tournade and Chabrol ¶ also showed that the secretion of adrenaline may be produced reflexly. Their method is to anastomose the suprarenal vein of one dog A, with the external jugular vein of a second dog B. The arterial pressure of B is recorded and the spleen of B placed in a plethysmograph. When the central end of the sciatic nerve of A was stimulated, there was a rise of blood pressure and a shrinkage of the spleen in B. By a similar technique, Heymans ** has shown that a fall of blood pressure in the arteries (but especially in the *sinus caroticus*) leads to a reflex liberation of adrenaline. The arrangement of this experiment is shown in Fig. 640. Adrenaline is liberated from the suprarenals of dog B when the carotid pressure on the left (innervated) side is lowered, but not when the carotid artery pressure on the right side is lowered (carotid sinus denervated on right side of B). When the adrenaline is liberated in B, the spleen of C (suprarenals removed) contracts.

Stewart and Rogoff,†† using other methods, claimed to have demonstrated that adrenaline is continually being passed out from the suprarenals into the blood, but it is doubtful whether in normal quietly resting animals the concentration of adrenaline in the blood exceeds 1 part in 1,000 millions, and this would produce no general effects.

When adrenaline is injected into the blood stream, the effect is only transitory—a matter of seconds for a small dose, *e.g.* 0·01 mg. for a cat. It rapidly disappears from the blood, and is taken up and protected from auto-oxidation by the tissues.‡‡ Here it seems to be inactivated in two ways, first, by an adrenaline oxidase, a variety of amine oxidase,§§ and secondly, (at all events when given by mouth) it is eliminated by conversion into an ester.* It is partly on this account that when injected into the portal vein adrenaline is relatively ineffective, since the liver is known readily to form sulphate esters with phenolic bodies.

---

* Anrep. *J. Physiol.*, 1912, **45**, 307.

† Kellaway. *J. Physiol.*, 1919, **53**, 211.

‡ Carrasco-Formiguera. *Amer. J. Physiol.*, 1922, **61**, 254.

§ Feldberg *et al.* *J. Physiol.*, 1934, **81**, 286.

‖ Cannon *et al.* *Amer. J. Physiol.*, 1921, **58**, 308, 338; 1924, **69**, 46.

¶ Tournade and Chabrol. *C. R. Soc. biol.*, 1926, **94**, 1199.

** Heymans. *Arch. intern. pharmacodyn.*, 1929, **35**, 269.

†† Stewart and Rogoff. *Amer. J. Physiol.*, 1917, **44**, 543; 1923, **66**, 235.

‡‡ Wiltshire. *J. Physiol.*, 1931, **72**, 88.

§§ Blaschko *et al.* *J. Physiol.*, 1937, **89**, 39P.

Adrenaline is thus a typical hormone, a body of comparatively low molecular weight, having a drug-like excitatory action on specific tissues of the body, easily diffusible, and rapidly destroyed after discharging its office. There seems little doubt that we must ascribe to the chromaffine tissue the function of elaborating this substance, and it has been claimed † that living medullary tissue can convert *p*-hydroxyphenylethylamine into adrenaline.

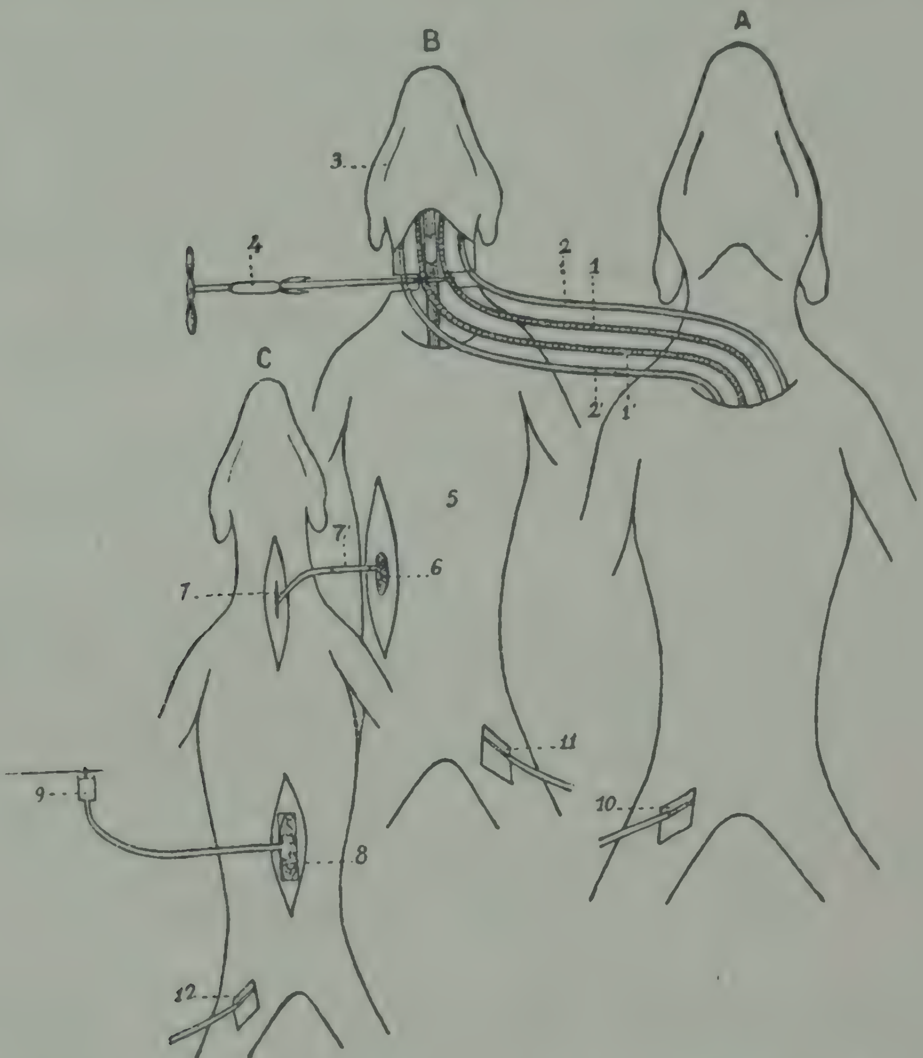


FIG. 640. Perfusion by Carotid and Jugular Anastomosis from A to Semi-isolated Head of B (right carotid sinus denervated), and from Suprarenal Vein of B to Jugular Vein of C (adrenalectomised). Spleen of C in a plethysmograph contracts when adrenaline is liberated by B, as *e.g.* when pressure in left carotid of B is lowered. Systemic pressures recorded from femoral arteries. Vagi of B divided; also all neck tissues, except spinal cord, divided by *écraseur* (4). (C. HEYMANS.)

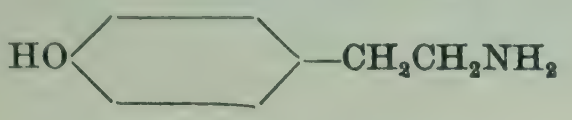
What part adrenaline plays in the normal functions of the body has, however, been contested with no little heat. It has actions closely resembling, but not identical with, those of the normal sympathetic transmitter (*nor*-adrenaline). An important difference is that, at low concentrations, it dilates muscle vessels, whereas *nor*-adrenaline constricts. The bulk of opinion seems to be that adrenaline is very slowly poured out by the normal suprarenal, and that this output is stopped or greatly reduced by section of the splanchnic nerve supply; that the amount of adrenaline so liberated is too small to exert an effect upon the tone of the arterioles, though it might affect the tone and permeability of the capillaries; and that, in states of emotion, asphyxia, exposure to cold, etc., there is a sudden liberation of adrenaline in such amount (*e.g.* 0·004 mg./kg./min.) as suffices to produce the effects seen on injection of that substance into the blood stream. According to this view, the *emergency theory*, we should be inclined to regard the suprarenal medulla as resembling the sympathetic nervous system in being a useful adjunct in times of stress, though not essential to life.

---

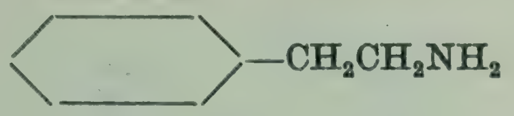
* RICHTER. *J. Physiol.*, 1940, **98**, 361.

† SCHULER and WIEDEMANN. *Z. f. p. C.*, 1935, **233**, 235.

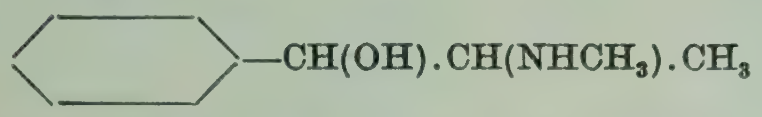
Owing to the rapid destruction of adrenaline, it produces very little effect when given by the mouth. There is, however, a whole series of substances, more or less allied to adrenaline in chemical constitution, which undergo less rapid destruction and can therefore be administered by mouth. Dale and Barger described such substances produced from amino-acids by a process of decarboxylation. Tyrosine would yield hydroxyphenylamine, and phenylalanine would give phenylethylamine. Such substances may be formed in minute quantities during the normal processes of putrefaction which occur in the alimentary canal. Another similar substance is the drug ephedrine. The constitution of these substances is shown in the following formulæ :



p-hydroxyphenylethylamine



phenylethylamine



ephedrine

Ephedrine sensitises tissues not only to adrenaline, but also to the stimulation of adrenergic nerves. This appears to be because it inhibits the enzyme, amine oxidase, by which adrenaline is destroyed.* This recalls the action of eserine in potentiating the action of acetyl choline or of cholinergic nerves.

For Reference

Britton (1930). Adrenal Insufficiency and Related Considerations. *Physiol. Rev.*, **10,** 617.

Broster *et al.* (1938). The Adrenal Cortex and Intersexuality. (London : Chapman and Hall.)

Grollman (1936). The Adrenals. (London : Baillière, Tindall and Cox.)

Loeb (1941). The Adrenal Cortex and Electrolyte Behavior. (*Harvey Lects.*, 1941–42, p. 100.)

Parkes (1945). The Adrenal-Gonad Relationship. *Physiol. Rev.*, **25,** 203.

Reichenstein (1943). The Hormones of the Adrenal Cortex. (New York : Acad. Press Inc.)

Swann (1940). The Pituitary-Adrenocortical Relationship. *Physiol. Rev.*, **20,** 493.

Swingle and Remington (1944). The Rôle of the Adrenal Cortex in Physiological Processes. *Physiol. Rev.*, **24,** 89.

THE THYROID

In development, the thyroid is formed by a down-growth from the floor of the pharynx. It consists of two oval bodies weighing, in man, 20 to 25 g., lying on either side of the trachea, and joined, in many animals, across the trachea by an isthmus. It is made up of an aggregation of vesicles varying in diameter from 15 to 150μ, each lined by a single layer of cubical epithelial cells, and filled with a fluid known as " colloid " (Fig. 643, A). Mitochondria are sometimes very abundant in the cells, and these and the Golgi apparatus undergo definite changes under conditions which are believed to be associated with alterations in the secretory activity of the thyroid, *e.g.* an increase when the animal is exposed to cold. The internal secretion formed by the epithelium is shed into the colloid, and probably makes its way into the blood direct, or along the lymphatic efferents of the gland. The thyroid is richly supplied with blood by three arteries, and is surrounded with a plexus of veins lying immediately under the capsule. It receives at least three times its own weight of blood per minute. The nerve supply is from the cervical sympathetic.

Diseases due to both deficient and excessive activity of the thyroid are known, and clinical observations of these first gave the clue to the functions of the thyroid.

Thyroid Deficiency. The effects of thyroid deficiency differ according to whether the deficiency is manifested in infancy or in later life.

If the atrophy of the thyroid occurs in early life, during the period of growth, *e.g.* in young children, the growth of the skeleton practically ceases. The bones of the limbs may grow in thickness but not in length. There is early synostosis of the bones of the skull, and complete cessation of development of mental powers. Children so affected are called *Cretins.* They may live for many years, but when twenty-five or thirty still present a childish appearance (Fig. 641, c). Stunted, pot-bellied and ugly, they have the intelligence of children of four or five. They often present fatty tumours above each clavicle, and similar subcutaneous tumours of fat or loose fibrous tissue are found in cases of myxœdema in the adult. The basal metabolism is low, and sugar tolerance high.

---

* Gaddum and Kwiatkowski. *J. Physiol.*, 1938, **94,** 87.

Removal of the thyroid from tadpoles causes a comparable failure of proper development, so that metamorphosis is absent or delayed.

If care be taken to spare the parathyroids at the operation, removal of the thyroid from young animals results in the same retarded development and stunting of growth as are observed in children when there is deficiency of the thyroid.

A decline of thyroid function in later life results in a somewhat different picture. In this country, attention was called to the connection of a depressed condition of metabolism known as *myxœdema* with alteration or atrophy of the thyroid. A patient affected with myxœdema presents a gradually increasing blunting of his or her mental activities ; speech is slow, cerebration delayed. With this nervous defect are associated changes in the connective tissues, the subcutaneous connective tissue becoming thickened, so that the face and hands appear swollen and puffy, looking at first sight as if œdema were present. The swelling is, however, due to newly-formed connective tissue, and not to the presence of an excess of interstitial fluid in the tissues. The patient often has a yellow waxy appearance, with a patch of colour on the cheeks. The hair falls out, the pulse is slowed, and the temperature tends to be subnormal owing to the diminution of the general metabolism of the body. Determination of the patient's basal metabolic rate shows this to be definitely (20–30 per cent.) below the normal value, as calculated from weight and height.* The intake of food and the excretion of nitrogen and of calcium are diminished. When the thyroid is extirpated in adults the result is the production of typical myxœdema.

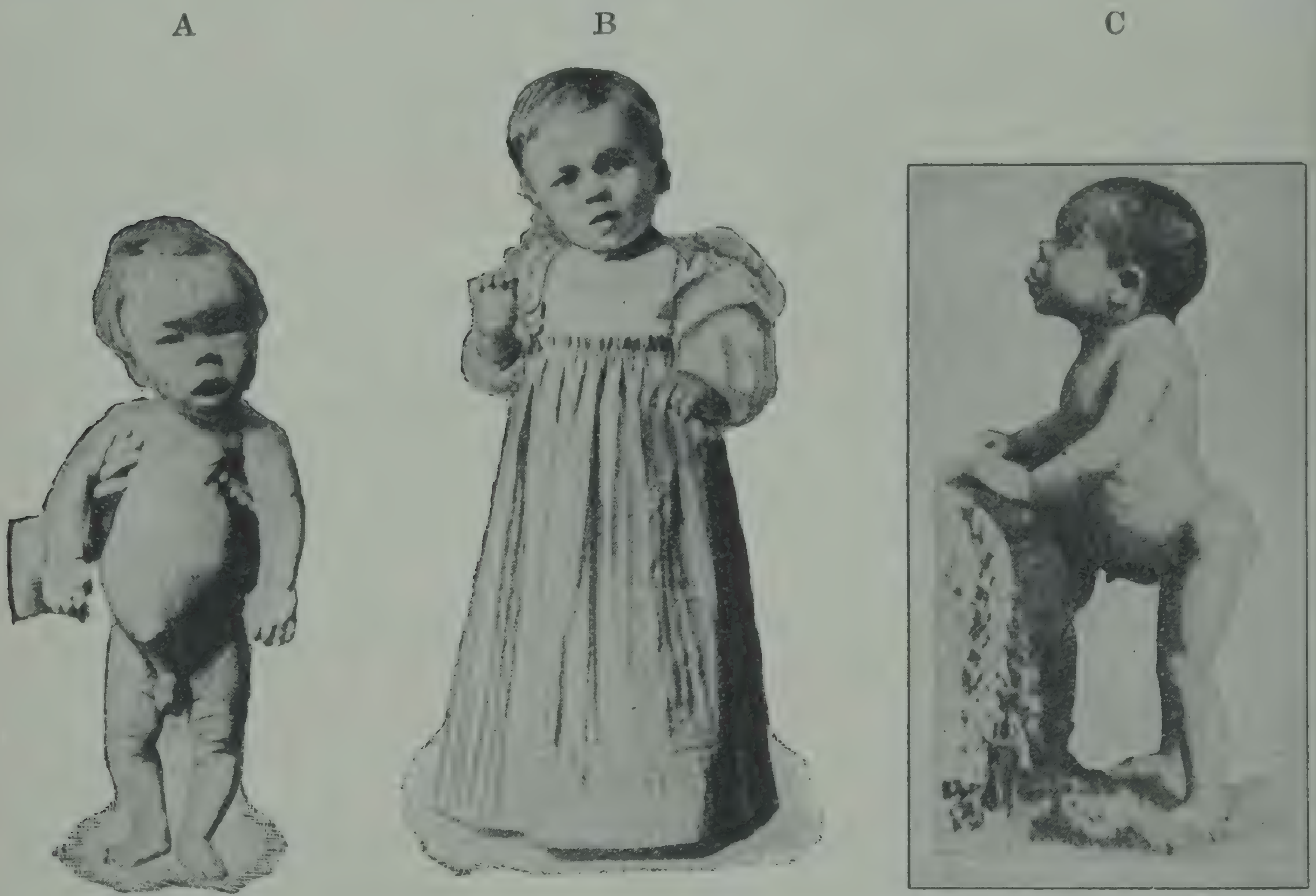


FIG. 641. A. A cretin, 23 months old. B. The same child, 34 months old, after administration of sheep's thyroids for 11 months. C. A cretin, untreated, 15 years old. (W. OSLER.)

THYROID FEEDING. Schiff first showed that the effects of extirpation of the thyroid might be prevented if, at the same time, the thyroid from another animal were transplanted into the subcutaneous tissue. On removing the transplanted thyroid, the typical symptoms of thyroid deprivation at once appeared. It was later found that it was sufficient to administer the thyroid gland, either fresh, dried, or partially cooked, by the mouth. An ugly and idiotic cretin can be converted by this means into a child of ordinary intelligence with normal powers of growth (Fig. 641, A and B). Given to myxœdematous patients, fresh or dried thyroid gland restores the basal metabolism, nitrogen output and body temperature to their normal levels, reduces the swelling of the subcutaneous tissues, causes a new growth of hair, and reinstates the patient's former condition of mental health *† (Fig. 642). It is possible that a moderate degree of thyroid inadequacy is not infrequent, and that the beneficial effects on general health, in removing excessive corpulence and in promoting the growth of hair, which are observed on administering thyroid to people in middle life, may be due to the actual reinforcement of a function which is being insufficiently discharged. Nor is the thyroid without influence on the healthy

---

* MAGNUS-LEVY. *Berl. klin. Woch.*, 1895, **32**, 650.

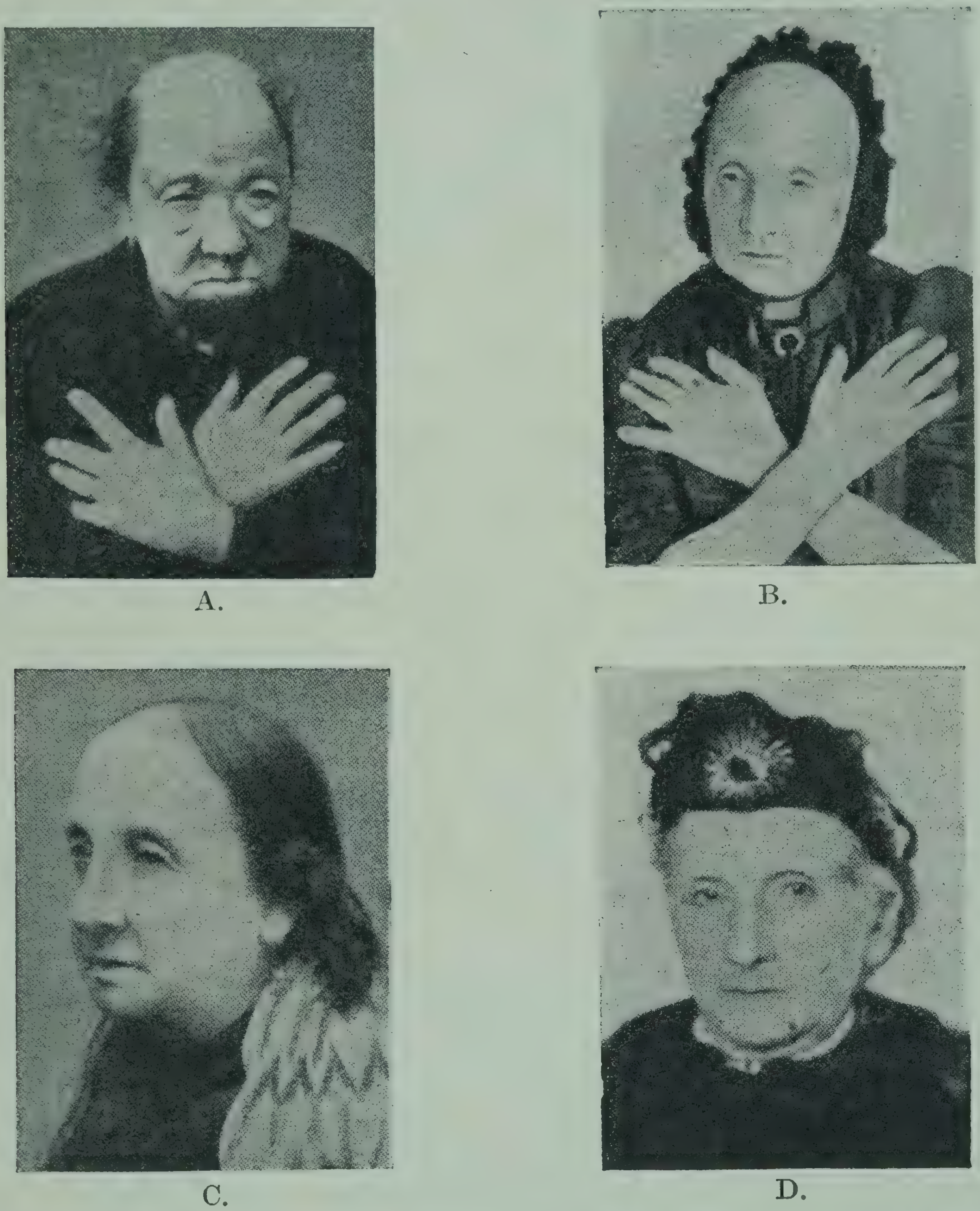


Fig. 642. Myxœdema treated for thirty years with Thyroid. A, aged 65, before treatment. B, five weeks later. C, fifteen months later. D, aged 94. (After Raven. *B.M.J.*, October 4th, 1924.)

---

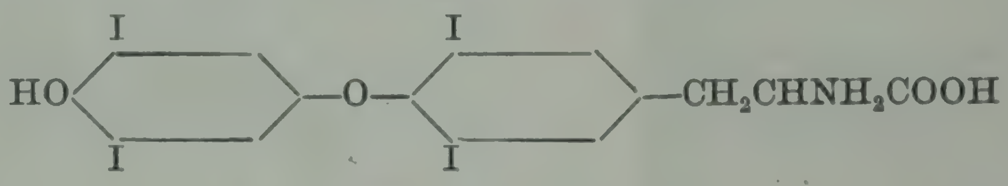
* Murray. *B.M.J.*, 1891 (ii.), 796.

† Fox. *B.M.J.*, 1892 (ii.), 941.

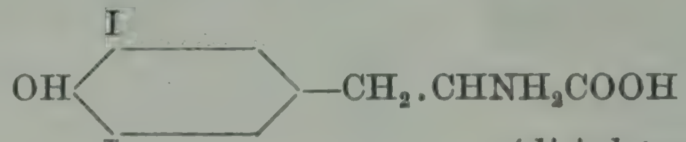
individual. When tadpoles are fed with thyroid substance, the rate of metamorphosis is greatly accelerated.* If given in large doses either to man or to animals, it quickens the pulse, even causing violent palpitation, and increases the metabolic activities of the body,† so that the appetite is increased, the glycogen store of the liver is reduced, the nitrogenous output rises above the intake, and the subcutaneous fat is diminished or disappears.

There are grounds for the belief that the thyroid plays a part in the production of the increased metabolism and rise of temperature in fevers. It may also, by liberating its secretion more rapidly, be concerned in causing the rise of metabolism which results when there is a tendency for the body temperature to be lowered ; under these conditions the blood flow through the thyroid is increased,‡ and the gland shows histological changes.

Thyroxine. The specific material of the colloid, called *thyroglobulin*, contains iodine in organic combination. Kendall, in 1914, isolated an active principle by hydrolysis of the thyroid, and gave it the name *thyroxine.* Harington and Barger § showed this to be the *l*-compound :



and prepared it synthetically. The *d*-isomer has one-sixth the activity of the natural one. About 60 per cent. of the iodine of the thyroid is present in another and simpler compound, *di-iodotyrosine:—*



(*di-iodotyrosine*)

which is physiologically much less active. Traces of mono-iodotyrosine also occur.

From two molecules of the di-iodo compound, by an oxidative reaction in which the side chain of one is removed, a molecule of thyroxine is formed. A similar reaction takes place when a tyrosine-containing protein is iodised, and by treating casein with iodine it has been shown that thyroxine can be obtained in small amounts.‖

The thyroid of an average human adult probably contains about 20 mg. of thyroxine. This substance, when administered, exhibits all the therapeutic properties of the dried gland itself. Thus 1 mg. of thyroxine definitely raises the basal metabolic rate, the maximal effect appearing after a day or two. It therefore increases the sensitivity of animals to oxygen lack. It also relieves the condition of hypothyroidism, and accelerates the metamorphosis of tadpoles. It is believed that the thyroid of the normal individual produces about 1 mg. thyroxine a day.

The thyroxine is manufactured in the thyroid itself, presumably from tyrosine, *viâ* iodotyrosine. It seems probable that the thyroxine is formed, and is stored, as part of a protein molecule, such as thyroglobulin. The gland has a remarkable affinity for inorganic iodine ; usually it stores about 2 to 5 mg. per g. of dry gland. If iodide be injected into the circulation, the greater part of it may be recovered from the thyroid within five minutes. The maximum uptake is about 10 to 20 mg./100 g. tissue.¶

Experimental evidence as to the biosynthesis of thyroxine has been obtained by the use of radioactive iodine (^{131}I). An injection of radioactive iodide is given to an animal, and within a few hours the iodine is found to have been concentrated in the thyroid. At first it is present as iodide : then some of it is found in the di-iodotyrosine fraction : later still, increasing quantities are in the form of thyroxine. Thus, in the sheep, the following amounts were found, as percentages of the total radioactive dose * :—

---

* GUDERNATSCH. *Amer. J. Anat.*, 1914, **15**, 431.

† ANDERSON and BERGMANN. *Skand. Arch. f. Physiol.*, 1898, **8**, 326.

‡ DIETRICH and SCHWIEGK. *Arch. f. exp. Path. u. Pharm.*, 1932, **165**, 53.

§ HARINGTON and BARGER. *Bioch. Journ.*, 1927, **21**, 169.

‖ HARINGTON and PITT-RIVERS. *Nature*, 1939, **144**, 205.

¶ LEBLOND and SÜE. *Amer. J. Physiol.*, 1941, **134**, 549.

Time.	As Di-iodotyrosine.	As Thyroxine.
2 hours	2–10 per cent.	0·1–1·0 per cent.
48 ,,	20–30 ,,	2·0–3·5 ,,

A similar demonstration has been effected with slices of surviving thyroid tissue : if the tissue is minced, there is no formation of thyroxine, nor is there if the tissue oxidations are inhibited by the addition of cyanide.† The essential reaction in the biosynthesis is probably an oxidation, whereby iodine is formed from iodide : the iodine converts tyrosine to di-iodotyrosine and this, while still in peptide combination, passes on to thyroxine spontaneously.

It is possible that the power of formation of thyroxine is not absolutely specific to the thyroid, since evidence of its formation in *small* amount has been obtained in the liver and muscle of thyroidectomised rats dosed with radioactive iodine. In the normal animal, however, the activity of the thyroid doubtless far outweighs that of the other tissues, which thus get but slight access to available iodine.

Although thyroxine produces effects so similar to those of the gland itself, the evidence that thyroxine, as such, is eliminated by the thyroid, or circulates in the blood is equivocal. Normal blood, however, does contain small amounts of iodine, about 3 μg. per 100 c.c., and this may be increased from two- to five-fold in exophthalmic goitre. By the intravenous injection of iodine, the iodine content of the blood is temporarily raised, but falls to normal in an hour or two, most rapidly in hyperthyroid subjects and most slowly in the hypothyroid ones.‡ The delay which elapses between the administration of thyroxine, and the production of physiological effects, has often been interpreted to mean that the relatively simple molecule of that substance needs to be built up into some more elaborate molecule before it can be effective. The active material of the thyroid is probably a globulin, of which thyroxine forms one of the constituents, and from which the active ingredient, thyroxine, may be liberated into the blood stream.

Thyroxine produces no effects when added to isolated tissues, although tissues taken from animals previously dosed with the active principles show an increased rate of oxygen usage. Thyroglobulin does cause an increase in the oxygen usage of isolated tissues,§ from which it was inferred that the delay in the action of thyroxine is due to the time needed for its conversion into thyroglobulin.

What causes release into the blood of the active constituent of the thyroid, whatever it may be, is not known. Some are in favour of a nervous influence, *viâ* the cervical sympathetic ‖ ; others regard the release as dependent on the presence in the blood of the thyrotropic pituitary hormone.

Iodine Lack. If iodine be entirely absent from the drinking water and food, the thyroid undergoes enlargement ; this is most clearly seen if the deprivation occurs during ante-natal life by reason of the mother before and during pregnancy having been fed on foods lacking in iodine. The resulting enlarged thyroid of the offspring has the structure of a parenchymatous hyperplastic goitre. The glandular tissue is increased in amount and the colloid reduced or absent (Fig. 643 B). If the lack of iodine is continued after birth the gland undergoes further hyperplasia. Early administration of small amounts of iodine in the form of iodide restores the gland to its normal size and structure, but if given later on in life complete restoration is unlikely, the gland instead undergoing further enlargement and showing on histological examination (Fig. 643, D) large irregular vesicles distended with colloid—the "colloid goitre." * Very large doses, however, may be more effective. Of common foods, fish is richest in iodine. Absence of iodine from the soil and water seems to be the cause of the great prevalence of goitre in certain inland districts, especially in Switzerland and in parts of the United States. It has been shown by Marine and others that this endemic goitre can be practically eliminated from these districts by the occasional administration of minute doses of iodine or iodides to the entire population (as by addition to the water supply or to the table salt).

---

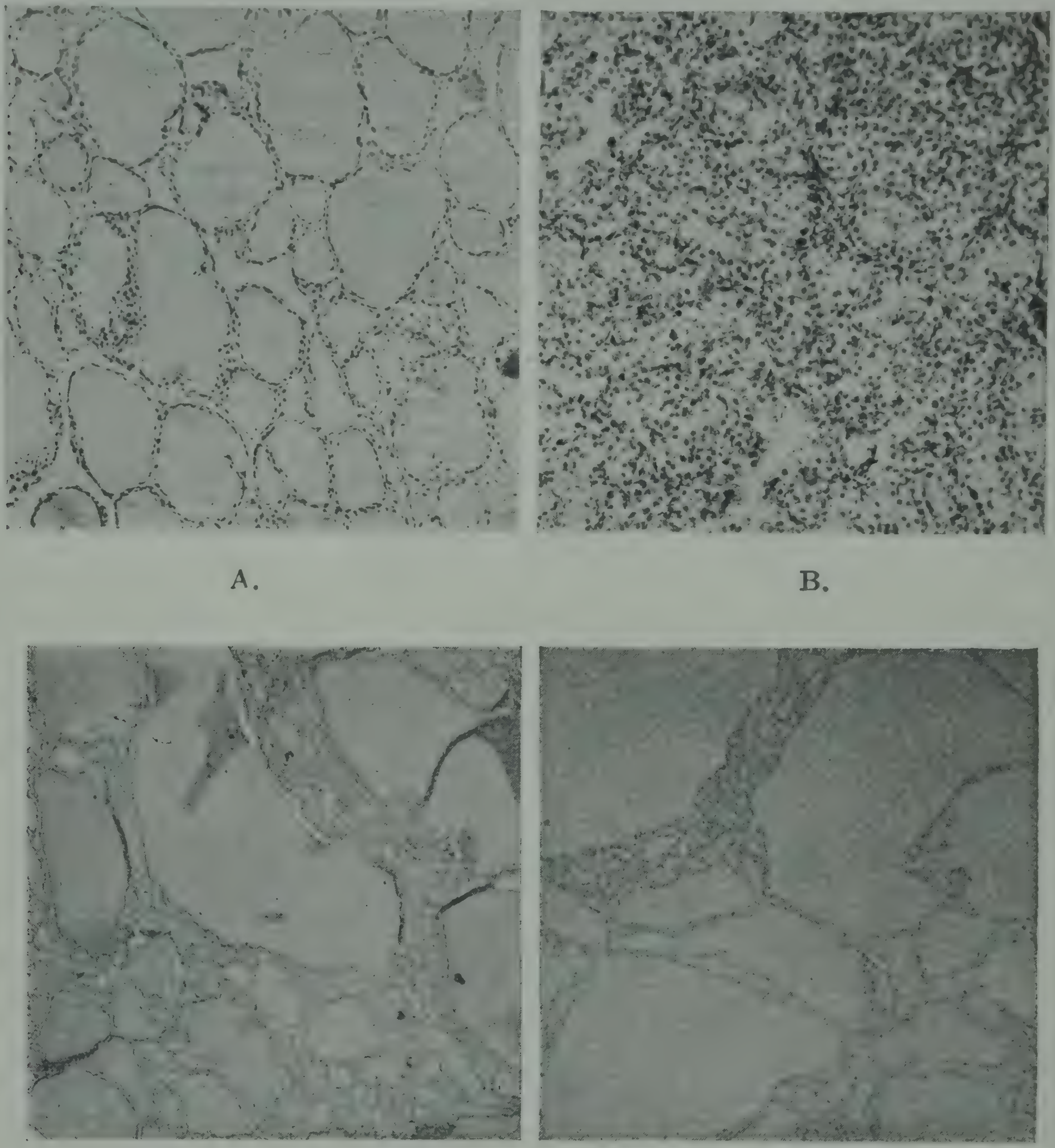
* Perlman, Morton and Chaikoff. *J. Biol. Chem.*, 1941, **139**, 449.

† Schachner, Franklin and Chaikoff. *J. Biol. Chem.*, 1943, **151**, 191.

‡ Watson. *Endocrinol*, 1938, **22**, 528.

§ Canzanelli *et al.* *Endocrinol.*, 1939, **25**, 707.

‖ Friedgood and Cannon. *Endocrinol.*, 1940, **26**, 142.



A. B.

C. D.

FIG. 643. Illustrating the Influence of Iodine Lack on Thyroid Structure. A = thyroid of puppy on diet which included cod-liver oil (containing iodine)—normal structure. B = thyroid of puppy on same diet except that it contained peanut oil (no iodine)—hyperplasia. C = human thyroid; simple colloid goitre. D = simple colloid goitre in dog's thyroid: the dog was iodine-deficient for 10 months, after which KI was given daily for 8 months (15μ g. I. daily.) (E. MELLANBY. "Nutrition and Disease." Oliver and Boyd, 1934.)

Hyperthyroidism. The symptoms caused by excessive doses of thyroid gland are strikingly similar to those occurring in the disease known as *exophthalmic goitre* (Fig. 644), where there is a true hypertrophy of the gland associated with increased basal metabolic rate, a body temperature tending to be above the normal, cardiac palpitation, exophthalmos or proptosis (bulging of the eyes), wasting, great restlessness, occasional glycosuria, and muscular weakness. Requirements of vitamins A, B and C are increased, because of the raised general metabolism. The iodine content of the blood is usually raised, while the thyroid itself contains less iodine than normally. Its structure is of the hyperplastic type, and this is possibly a result of its low iodine content. Administration of iodine causes a transitory return of the structure towards the normal, with some reduction in the symptoms, but these become worse later on, especially if the iodine is discontinued.

---

* E. MELLANBY. Nutrition and Disease. (1934: Oliver and Boyd.)

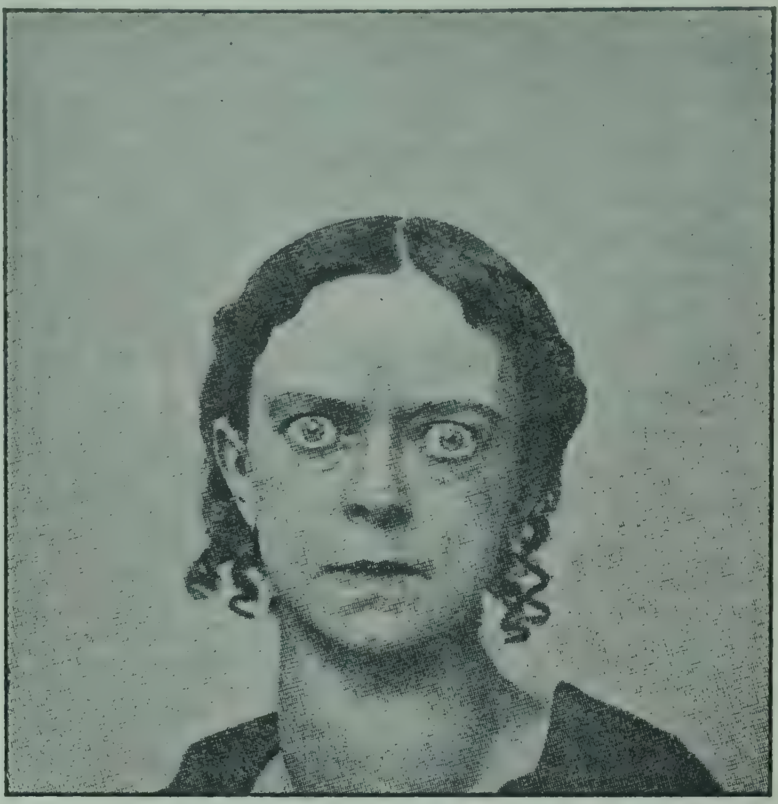


FIG. 644. Exophthalmic Goitre. (After PARSONS.)

In these cases, removal of part of the hypertrophied gland often removes the symptoms and restores the patient to health. In some cases, however, the condition recurs. If the whole gland is removed, it is necessary to administer daily doses of dried gland for the rest of the patient's life, in order to prevent the development of myxœdema.

One view of the condition is that it is due to a failure of the gland to retain the thyroxine which it makes. This is attributed to a hormone present in the anterior lobe of the pituitary body.

Thyrotropic Hormone. As will be shown later, the pituitary anterior lobe is closely related to the activity of the thyroid. Extracts of the anterior pituitary can be prepared which, on injection, cause hypertrophy and hyperactivity of the thyroid, and even exophthalmos. Indeed, exophthalmos is believed to be directly due to the presence of the thyrotropic hormone and not to excess of thyroxine. Injection of such extracts causes increase in basal metabolism, hyperplastic change of structure in the thyroid, and increase of blood iodine.*

Thyroid and Vitamins. The condition of hyperthyroidism, or the administration of thyroid, or of thyroxine, increases the requirement of vitamin B_1, lack of which has been demonstrated in the tissues. This is because the B_1 requirement varies with the calorie output, which thyroxine raises. If adequate amounts of the vitamin are given, the loss of weight accompanying thyroid administration is lessened.† The effects of vitamins A ‡ and C is similar, as regards loss of weight.

Goitrogenic Substances. Administration of certain compounds by mouth leads to enlargement of the thyroid, with histological changes similar to those seen in parenchymatous goitre. Such substances include seeds of *Brassica*, KCNS, thiourea, thiouracil, allylthiourea, *p*-aminobenzoic acid, and sulphaguanidine. Administration of 0·1 per cent. thiouracil in the drinking water of rats, for example, soon produces goitres, and reduces the B.M.R. This substance has been used clinically, with striking results, in the treatment of hyperthyroidism.

---

* KROGH, LINDBERG and OKKELS, SCHITTENHELM and EISLER. *Cit.* E. Mellanby.

† PETERS and ROSSITER. *Bioch. J.*, 1939, **33,** 1140.

‡ LOGARAS and DRUMMOND. *Bioch. J.*, 1938, **32,** 964.

The effect of the thiouracil is to depress the synthesis of thyroxine in the thyroid, probably by opposing the oxidative conversion of iodide to iodine: it does not abolish the effects of a dose of thyroxine: KCNS on the other hand inhibits the uptake of iodine by the thyroid.*

The B.M.R. of rats is also lowered by daily dosage of 1–7-dimethylxanthine (paraxanthine), about 10 μ.g./100 g. body weight. The substance, which is of widespread occurrence, appears to antagonise the effects of thyroxine.†

For Reference

Harington (1933). The Thyroid Gland. (London: Oxford University Press.)

Harington (1944). Thyroxine: its biosynthesis and its immunochemistry. *Proc. Roy. Soc.*, **132B**, 223.

Drill (1943). Interrelations between Thyroid Functions and Vitamin Metabolism. *Physiol. Rev.*, **23**, 355.

McClendon (1927). The Distribution of Iodine with Special Reference to Goiter. *Physiol. Rev.*, **7**, 189.

Marine (1922). The Present Status of the Functions of the Thyroid Gland. *Physiol. Rev.*, **2**, 521. Marine (1935). The Thyroid. *J. Amer. Med. Assoc.*, **104**, 2250.

Salter (1940). Fluctuations in Body Iodine. *Physiol. Rev.*, **20**, 345.

THE PARATHYROIDS

The parathyroids, discovered by Sandström in 1880,‡ are four oval bodies, which, in man, together weigh about 140 mg., situated, two on each side, on the border of the thyroid gland or actually embedded in its substance. One of each pair is developed from the 3rd, and one from the 4th branchial pouch. In histological appearance, they consist of solid masses of columns of epithelial cells, most of which are pale and hyaline, but a few of which are oxyphil, surrounded with connective tissue and richly supplied with blood sinusoids (Fig. 645). In some animals, *e.g.* in the dog, they are embedded in the thyroid. In others, such as the rabbit, they lie outside the gland, and it is easy to avoid them in the exision of the thyroid. Accessory parathyroid tissue is also frequently found in the thymus and along the aortic arch. To this varying distribution of the parathyroids is due the discrepant early results of extirpation of the thyroid in carnivora and herbivora respectively. The parathyroids have often been inadvertently removed in man, with fatal results, in the operation of thyroidectomy, prior to a knowledge of their importance.

In spite of their close anatomical association with the thyroid, the parathyroids have a function entirely distinct from that gland. Whereas removal of the thyroid produces simply a condition of myxœdema, or retardation of growth in the young animal, removal of the parathyroids is followed by nervous disturbances and *tetany*, which rapidly prove fatal, especially in young animals.§ The motor nerves become inordinately excitable to galvanic currents. In man, tetany consists in a tendency for muscles, especially of the forearms, to undergo painful spasms. There is flexion of the metacarpo-phalangeal joints, the interphalangeal joints being extended, and thumb and little finger are drawn towards the palm. Wrist and elbow may be flexed and forearm slightly pronated. The carpal spasm may be precipitated by pressure on the upper arm with a sphygmomanometer armlet (Trousseau sign). The condition is always found to be accompanied by a diminution of the calcium content and increase of P content of the blood,‖ the former falling from the normal figure of about 10 mg. Ca per 100 c.c. serum to between 4 and 7 mg. The excretion of calcium and of phosphate is reduced. In chronic cases of hypoparathyroidism defects of skin. hair and nails occur, and cataract of the lens is encountered.

---

* Franklin, Chaikoff and Lerner. *J. Biol. Chem.*, 1944, **153**, 151.

† Carter *et al.* *Nature*, 1943, **151**, 728.

‡ Upsala Läk. Förh. 1880, **15**, 441. [English translation by Seipel, *Bull. Inst. Hist. Med.*, 1938, **6**, 179.]

§ Gley. *C. R. Soc. Biol.*, 1897, **4**, 18.

‖ MacCallum and Voegtlin. *J. Exp. Med.*, 1909, **11**, 118.

The symptoms are found to be temporarily relieved by administration of calcium lactate, either by the mouth or subcutaneously, and are aggravated by a meat diet. The condition of tetany after parathyroidectomy may be relieved by subcutaneous implantation of a normal gland.

Collip * discovered that, by extraction of fresh parathyroid glands with 5 per cent. hydrochloric acid and then precipitating the proteins by neutralisation, a fluid is obtained which contains the hormone of the parathyroids in a very potent form. One or two injections per day of this extract serves to ward off indefinitely the symptoms of tetany in parathyroidectomised dogs, or to cure such symptoms when they have already supervened, either in dogs or in human beings suffering from tetany due to parathyroid insufficiency. The parathyroids are thus the seat of formation of a hormone which regulates calcium metabolism through a specific control of blood calcium concentration.

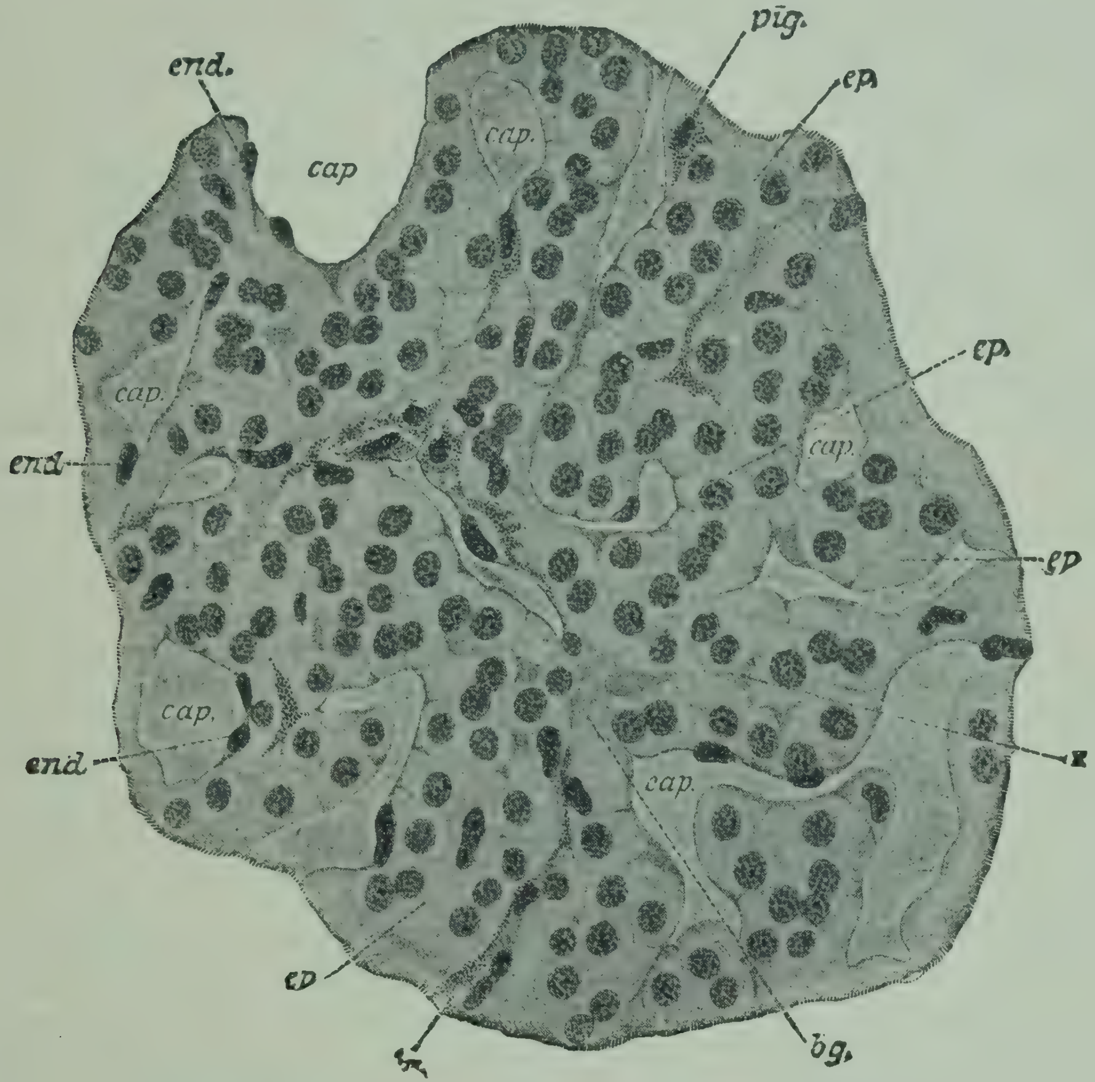


FIG. 645. Section of Parathyroid. (KOHN.)

ep.	Secreting epithelium.	*cap.*	Sinus-like capillaries.
pig.	Cells containing pigment.	*end.*	Endothelial cells.

The active principle is a protein,† and is destroyed by pepsin, so that it is ineffective when given by the mouth.

When the extract is administered to normal dogs there is, after a lag of three to four hours, a steady rise in the calcium content of the serum, which reaches a maximum in five to nine hours, and then returns at the same rate to normal. The blood phosphorus is lowered as the calcium rises, except when large doses are given. Excretion of calcium and phosphorus are increased. When several successive injections are made, there is a summation of effect, and the calcium value of the serum may rise to 15 or 20 milligrammes per 100 c.c. When the figure of 15 is passed, the dog shows definite symptoms of "hypercalcæmia." The animals lose appetite, become weak, there is diarrhœa and vomiting and rise of blood urea, followed by collapse and death. The impairment of the circulation is associated with great loss of plasma, and increase in the viscosity of the blood. Prolonged overdosage may lead to the formation of renal calculi, which consist of calcium phosphate, and to calcifications of various tissues.

Parathyroid extracts ("parathormone") are standardised by administration to dogs and determination of the increase in serum calcium in a given time. The unit is defined as 1/500th of the amount required to produce in a dog of 20 kg. weight an increase of 5 mg./100 c.c. in serum calcium in fifteen hours.* The most potent dry preparations contain about 12 units per mg. In man 100 units per day will raise a low blood calcium by 50 per cent., but when the blood calcium is normal even 500 units may only cause a small rise.

---

* COLLIP. *J. Biol. Chem.*, 1925, **63**, 395; *Amer. J. Physiol.*, 1926, **76**, 472.

† ROSS and WOOD. *J. Biol. Chem.*, 1943, **146**, 49.

Calcium of Blood and Bone. The action of the parathyroid hormone is closely connected with the complex and largely unsolved problems of calcium metabolism. Of the normal serum content of calcium of about 10 mg./100 c.c., probably less than half is present in the ionizable form, the remainder being in the form of non-diffusible and unionised complexes of uncertain nature. The ratio between these two forms seems to depend on the *pH*, the protein, the phosphate, and probably other things. Thus, if phosphate is added to the blood, either *in vitro* or *in vivo*, there is a conversion of Ca ions into a physiologically inert and colloidal form, as calcium phosphate.† In the body this colloidal calcium phosphate is phagocytosed by macrophages of the liver and spleen, and so removed from circulation.‡ The excitability of the neuro-muscular mechanisms appears to depend chiefly on the content of ionizable calcium present, and the causation of tetany by rise of *pH* (as by overbreathing) may similarly be ascribed to reduction of the ionic calcium of the blood and tissues.

The principal store of calcium is in the bones. The nature of the mineral basis of bone is still a matter for dispute; it contains calcium phosphate and carbonate, or hydroxide, in some close chemical or physico-chemical union, and appears to be of constant composition.§ The relations of $Ca^{\cdot\cdot}$ and PO_4''' ions to calcium metabolism have been discussed briefly on p. 817, and it was there stated that the calcification of bone is probably dependent upon the liberation of phosphate from organic combination owing to the action of the enzyme phosphatase, which is present in the bone. In this and in many other ways the metabolism of calcium and of phosphate are closely related.

Under the influence of calcium lack, or of the action of the parathyroid hormone, calcium is removed from the bones, especially from the trabeculæ of the cancellous bone. The excretion of calcium and of phosphate in the urine is increased. In lead poisoning, lead is deposited with calcium in the bone, and may remain there for long periods if the calcium is not mobilised; if a dose of parathyroid hormone is given, however, the lead is released with the calcium,‖ and excreted. If a period of rest is now given, the calcium is replaced in the bone, and, on giving a second dose of parathyroid, calcium without lead is excreted. It seems that the parathyroid acts by stimulating the formation of new osteoclasts, which exert their solvent action on the calcium salts of the bone. At a later stage of overdosage, however, a reaction may occur, and the animal becomes relatively immune: in this stage osteoblasts are increased and new bone formed rapidly.¶ It has also been thought that the action of vitamin D may be due to a stimulant action on the parathyroid, but the evidence for this is not free from objections. The whole subject is complex, and the reader is referred to the reference articles for a fuller discussion.

Hyper- and Hypo-parathyroidism. Occasionally cases of parathyroid adenomata occur, and show all symptoms of hyper-parathyroidism—high Ca and low P in the serum, increased Ca and P excretion, softening or fibrocystic disease of the bones, and often renal phosphatic calculi. Removal of the tumour, which may weigh several grammes, cures the condition. The converse state of hypoparathyroidism, accompanied by low blood Ca, high P, reduced Ca and P excretion, and signs of tetany, may occur spontaneously, or after surgical operations on the thyroid.

For Reference

Dragstedt (1927). The Physiology of the Parathyroid Glands. *Physiol. Rev.*, **7**, 499.
Kay (1932). Phosphatase in Growth and Disease of Bone. *Physiol. Rev.*, **12**, 384.
Thomson and Collip (1932). The Parathyroid Glands. *Physiol. Rev.*, **12**, 309.
Stewart and Percival (1928). Calcium Metabolism. *Physiol. Rev.*, **8**, 283.

THE PITUITARY BODY (Hypophysis)

In the pituitary body, which weighs about 0·5 g. in man, we have four important structures, the *anterior lobe, posterior lobe, pars intermedia,* and the *pars tuberalis.*** It consists of two main parts, anterior and posterior, which have separate modes of origin. An outgrowth from the buccal cavity in the embryo meets a hollow extension of the anterior cerebral vesicle. The buccal ectoderm gives rise to the anterior lobe and pars intermedia of the pituitary, while the neural epiblast becomes developed into the posterior lobe (Fig. 646). In some animals, the posterior lobe remains hollow and retains its primitive connection with the third ventricle of the brain, but in man it becomes entirely solid.

---

* Collip's original unit was 1/100th of the amount so required. The fivefold reduction was agreed upon in 1934.

† McLean and Hinrichs. *Amer. J. Physiol.*, 1938, **121**, 580.

‡ Gersh. *Ibid.*, p. 589.

§ Kay. *Ann. Rev. Biochem.*, 1932, **1**, 189. Thompson and Collip. *Physiol. Rev.*, 1932, **12**, 340.

‖ Hunter. *Lancet*, 1930, **218**, 897, 947, 999.

¶ Selye. *Endocrinol.*, 1932, **16**, 547.

** Herring. *Q. J. Exp. Physiol.*, 1908, **1**, 121, 161, 261; 1914, **8**, 245, 267.

The anterior lobe in the adult, which comprises pars tuberalis, pars intermedia, and pars distalis, consists of nests of epithelial cells, many of which, the chromophobe cells, are only stained with difficulty, while some (the α-cells) are filled with acidophile and others (the β-cells) with basophile granules. The chromophobe cells are parent cells from which the two types of chromophile cells are formed. In man there are normally about 50 per cent. chromophobe, forty of α- and ten of β-cells.

Striking changes in the histology of the anterior lobe occur under certain conditions. For example, after castration the basophile cells are increased: in pregnancy, cells called pregnancy cells appear; they are finely granular acidophile cells, derived from chromophobe cells; after thyroidectomy there is diminution in the acidophiles and vacuolation of basophiles; in hyperthyroidism the acidophiles increase; œstrone injections lead to increase, and then to degeneration of the basophiles, later the acidophiles are increased.

The anterior lobe is separated from the posterior lobe by a cleft, which is the remains of the original hollow outgrowth from the buccal cavity. The anterior lobe and pars tuberalis are highly vascular, the other parts only poorly supplied with vessels. The epithelial tissue immediately adjoining the cleft probably belongs to the *pars intermedia* and differs from that constituting the anterior lobe. Its cells, which present but few granules, are arranged in islets, separated by an intervening tissue continuous with the main mass of the posterior lobe. Many of the islets are hollow and enclose a colloid material. The amount of colloid material increases in animals which have undergone extirpation of the thyroid gland.

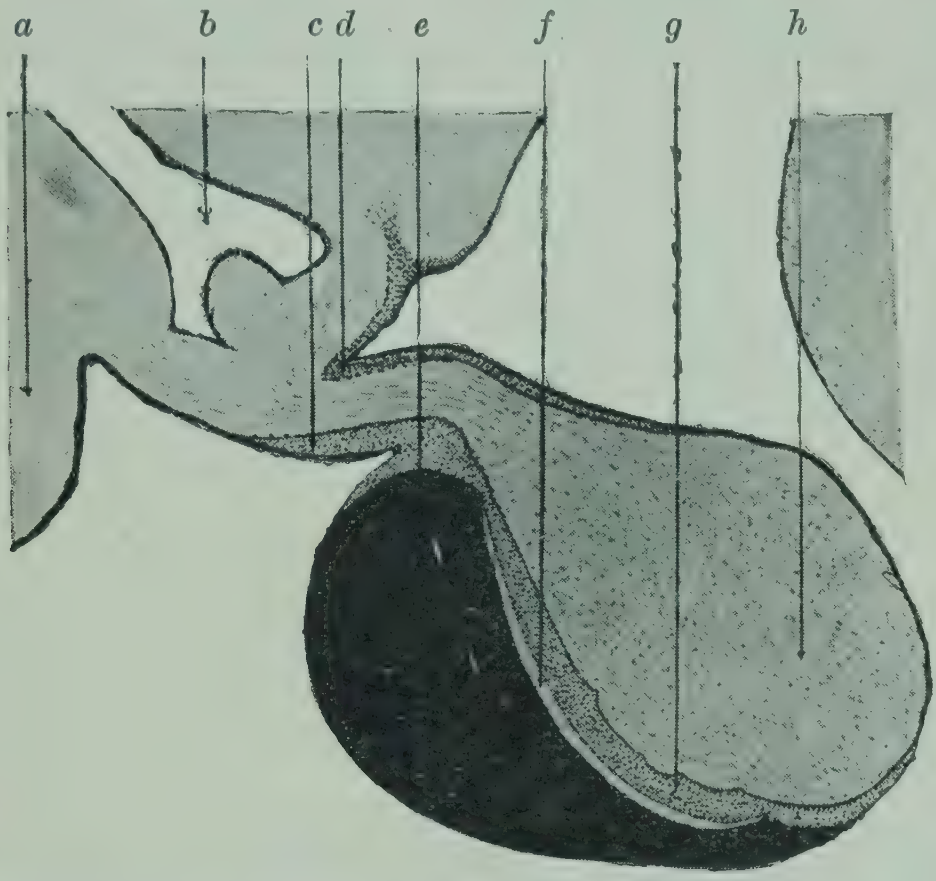


FIG. 646. Mesial Sagittal Section through the Pituitary Body of an Adult Monkey. (Semi-diagrammatic.) (After HERRING.)

a.	Optic chiasma.	*e.*	Anterior lobe.
b.	Third ventricle.	*f.*	Epithelial cleft.
c.	Pars tuberalis.	*g.*	Pars intermedia.
d.	Epithelial investment of posterior lobe.	*h.*	Posterior lobe.

The colloid material of the anterior lobe has been traced by Herring into the stalk and is believed to contain active materials which may ultimately find their way into the third ventricle. An interesting vascular arrangement has been described * under the name hypophysio-portal circulation. In this, arterial twigs supply the median eminence of the posterior pituitary, the efferent vessels from which then form wide sinusoids in the pars distalis of the anterior lobe. The anterior lobe appears to have no effector nerves, such few nerve fibres as it has being chiefly related to the blood vessels.† It has been suggested that its activities are controlled by humoral means, perhaps by way of the hypophysio-portal system.

The *posterior lobe* includes the median eminence of the tubular cinereum, the infundibular stem, and the infundibular process. The last-named part is composed of neuroglial tissue and spindle-shaped cells called *pituicytes.*‡ It is also abundantly supplied with nerve fibres, many of which terminate in structures that appear to be sensory receptors, and it has been suggested that these endings may "constitute a 'chemo-receptive' or 'presso-receptive' apparatus devoted to the reception of stimuli which enable the hypothalamic centres to control the metabolic and hormonal functions of the body."[*] The posterior pituitary receives sympathetic post-ganglionic nerve fibres; their cell stations are in the superior cervical ganglion. It is perhaps of importance that there are nervous connections (the supraoptico-hypophyseal, paraventriculo-hypophyseal, and the tubero-hypophyseal tracts, Fig. 656) between the hypothalamus and the pituitary. These are efferent tracts which can be traced to the posterior lobe.

---

* POPA and FIELDING. *J. Anat.*, 1933, **67**, 227; GREEN and HARRIS. *J. Endocrin.*, 1947, **5**, 136; HARRIS. *J. Anat.*, 1947, **81**, 343.

† RASMUSSEN. *Endocrin.*, 1938, **23**, 263.

‡ BUCY. *J. Comp. Neurol.*, 1930, **50**, 505.

The *pars tuberalis* surrounds the infundibular stalk, and spreads out on the base of the brain in the neighbourhood of the tuber cinereum. The interest of the pars tuberalis is due to the fact that it has been found that many of the results attributed to removal of the pituitary body can be obtained by merely injuring the hypothalamic region; these results are polyuria, glycosuria, loss of heat regulation, and early death.[†] An important feature of the results is atrophy of the posterior lobe.

The two lobes, anterior and posterior, certainly have entirely different functions. The *pars intermedia* is closely adherent to the posterior lobe, and its functions have not been entirely dissociated from those of the latter.

THE ANTERIOR LOBE

Extracts of the anterior lobe of the pituitary certainly contain at least six different hormone-like substances, and though it cannot be definitely asserted that these same substances are passed into the blood stream, they are provisionally assumed to be. They are named as shown below :—

Metabolic Hormones	(1)	(*)	Growth and diabetogenic hormone.
	(2)		Thyrotropic hormone.
	(3)	(*)	Adrenocorticotropic hormone (ACTH).
Gonadotropic Hormones	(4)		Follicle-stimulating hormone (FSH).
	(5)	(*)	Interstitial-cell-stimulating hormone (ICSH).
	(6)	(*)	Lactogenic hormone.

All these anterior pituitary hormones are proteins. Those marked * have been separated in the pure crystalline state. (1) and (6) are only slightly soluble in water, the others are freely soluble. The proteins are of two types, (1) (3) and (6) being simple proteins, and (2) (4) and (5) being glyco-proteins, containing mannose and hexoseamine. The proteins show species specificity in some cases.

Since these hormones have the property of stimulating other endocrine organs, it is evident that the anterior pituitary could, by releasing them, exercise profound influences over the whole body, and that the description of the anterior pituitary as the master gland of the endocrine system is correct cannot be doubted.

Effects of Extirpation. The effects of extirpation of the whole pituitary (hypophysectomy) are mainly due to the removal of the anterior lobe. Extirpation of this lobe, by abruptly removing the source of all the anterior pituitary hormones, as would be expected, produces extensive and complex results. In clinical cases of pituitary hypofunction in which different factors may be affected to different extents, and in which, accordingly, the symptoms are very variable, effects identical with those seen on experimental extirpation are not to be anticipated, and in fact are rarely, if ever, encountered. Our fundamental clues have, however, been gleaned from clinical observation, and confirmed and supplemented by properly controlled laboratory experiments on the lower animals.

The pituitary body is not essential to life, but its extirpation is followed by profound changes.[‡] Growth is arrested sooner or later, there is loss o

---

[*] Vazquez-Lopez. *Brain*, 1942, **65**, 1.
[†] Cushing. *Lancet*, 1930 (ii.), 119, 175.
[‡] McPhail. *Proc. Roy. Soc.*, 1935, **117B**, 45.

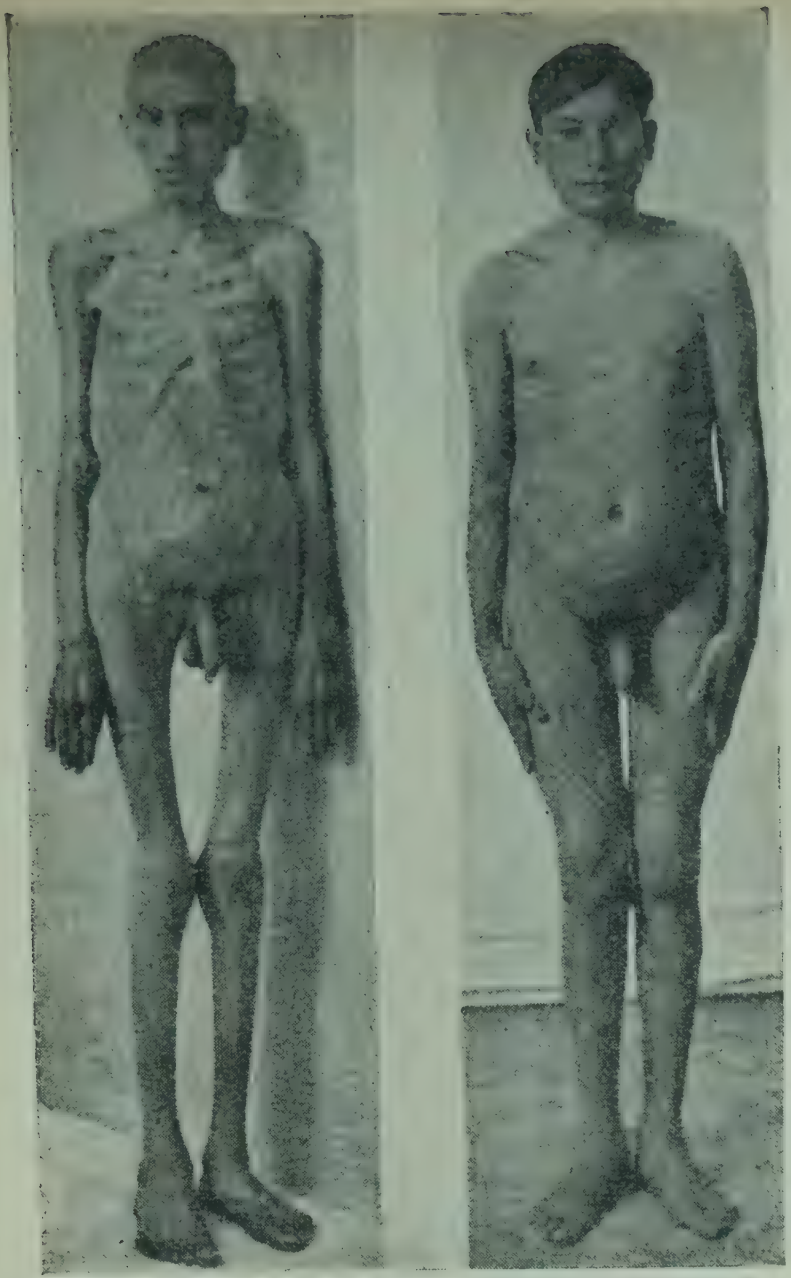


Fig. 647. Simmond's Disease before and after Treatment by Anterior Pituitary Injections. (From Lucacer. *La Riforma Med.*, 1932, **48**, 547.)

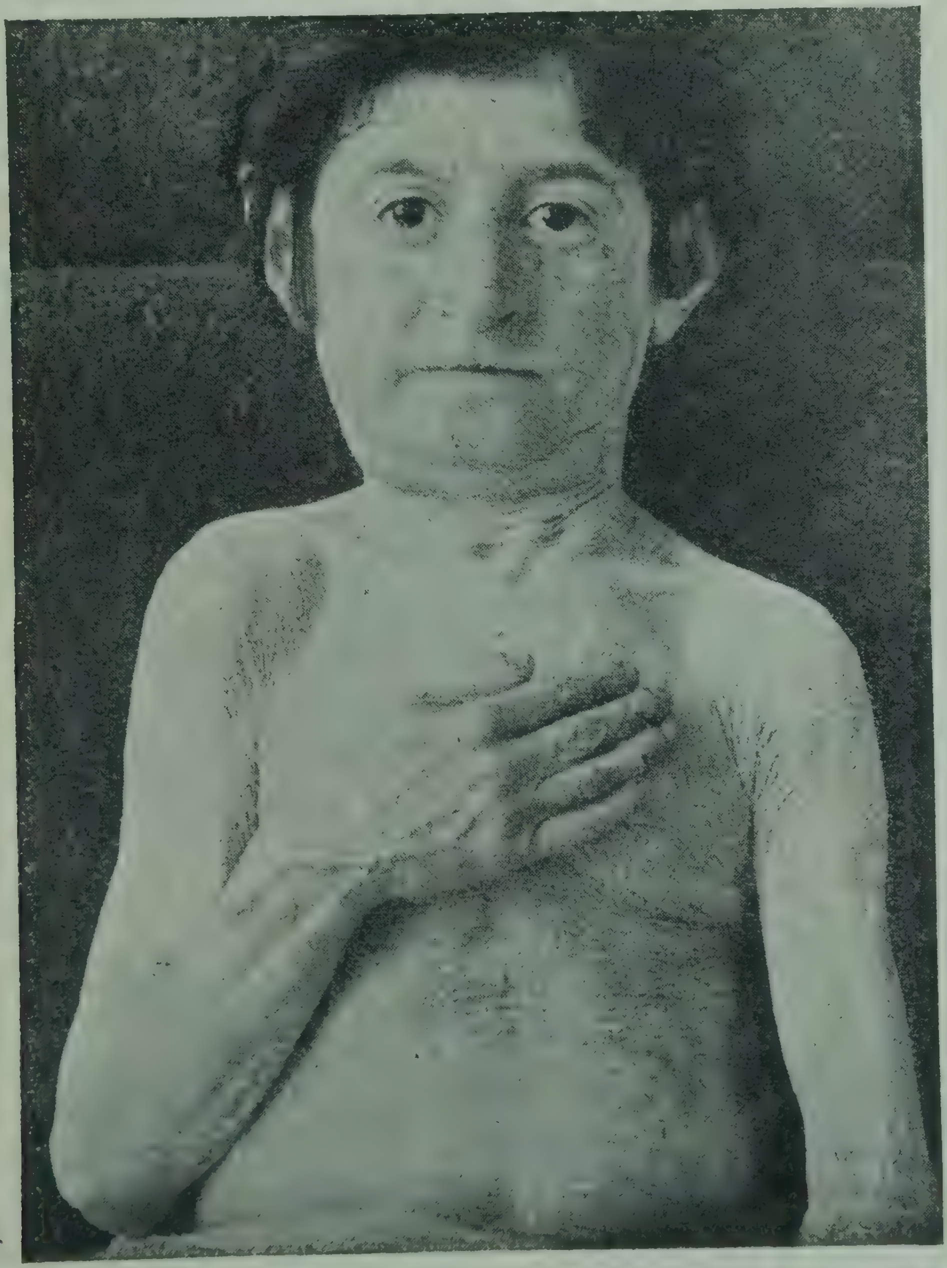


Fig. 648. Progeria in a Girl 8 years of age. (From a photograph kindly lent by Professor Harvey Cushing.)

muscle tone, the reproductive organs fail to mature, or else atrophy, there is atrophy of the thyroid, suprarenal and probably other endocrine organs and profound metabolic changes. There is a high sugar tolerance, a reduction in the hyperglycæmic action of adrenaline, and the post-absorptive blood sugar content is lowered. Protein metabolism in fasting is at first increased and later diminished: there is often abundant deposition of adipose tissue. This latter is, however, stated to be due rather to incidental injury to the hypothalamic region than to the removal of the pituitary itself.

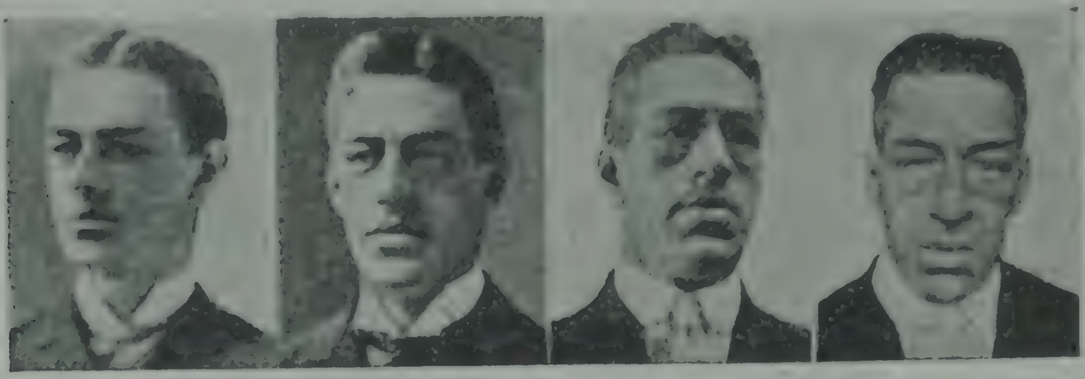


FIG. 649. A Case of Acromegaly. I., aged 24, before onset of the disease. II., aged 29. III., aged 37. IV., aged 42, with pronounced acromegaly. (After CUSHING. "The Pituitary Body and its Disorders." Lippincott, 1912.)

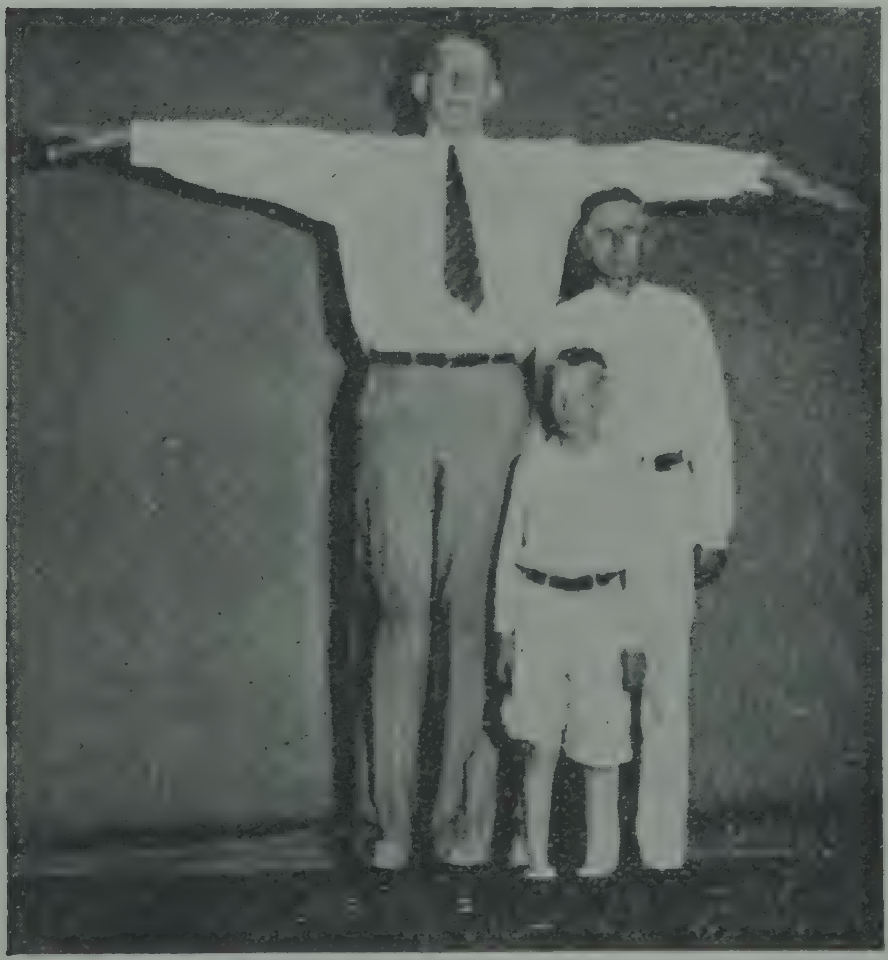


FIG. 650. Giant, aged 13½, beside a 9-year-old Brother, and his Father, who is 5 ft. 11 in. in height. (After BEHIEMS and BARR. *Endocrinology*, 1932, **16**, 125.)

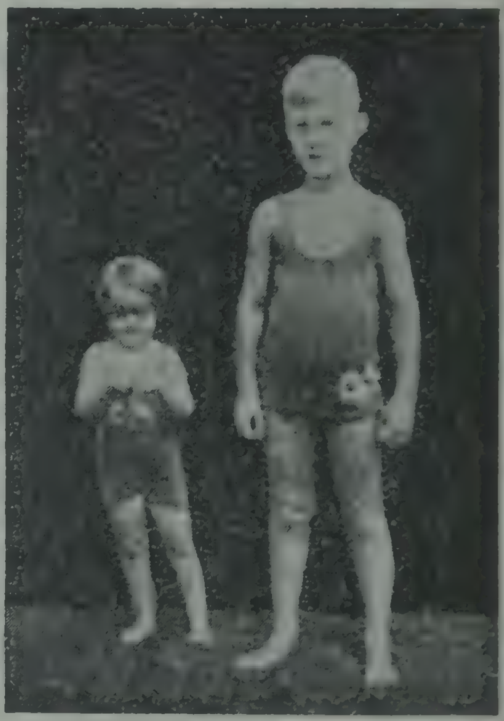


FIG. 651. Female Pituitary Dwarf aged 9½ compared with a Normal Boy of the same age. (After ENGELBACH. *Endocrinology*, 1932, **16**, 11.)

In the human subject, generalised hypofunction of the anterior lobe is found in Simmond's disease (Fig. 647); there is lowered metabolism, cachexia, loss of sexual function and premature senility. In the disease called progeria, which is also probably of pituitary origin (Fig. 648), the senescence is very dramatic. Atrophy of the anterior lobe is found at autopsy.

It is interesting to note that there is a reciprocal influence between the anterior pituitary and the various organs which it can affect. Thus, pregnancy, or extirpation of the gonads, thyroid or suprarenals causes definite histological changes to occur in the anterior lobe of the pituitary body.

(1) The Growth Hormone.* The first clue to the functions of the anterior pituitary body was furnished by the observations of Marie † that the disease called *acromegaly* is associated with pituitary tumour, and it is now known that the tumours consist of the eosinophile cells. In this disease certain parts of the skeleton, particularly of the lower jaw, hands and feet, take on renewed growth in adult life (Fig. 649). Various other symptoms not directly related to the increase in the production of growth hormone are also often present. Thus there may be limitation of the field of vision, leading to blindness; this begins in the temporal fields owing to pressure of the tumour on the optic nerve fibres from the nasal parts of the retinæ. The condition is relieved by surgical removal of the tumour. There is also not infrequently a lowered sugar tolerance or even diabetes.

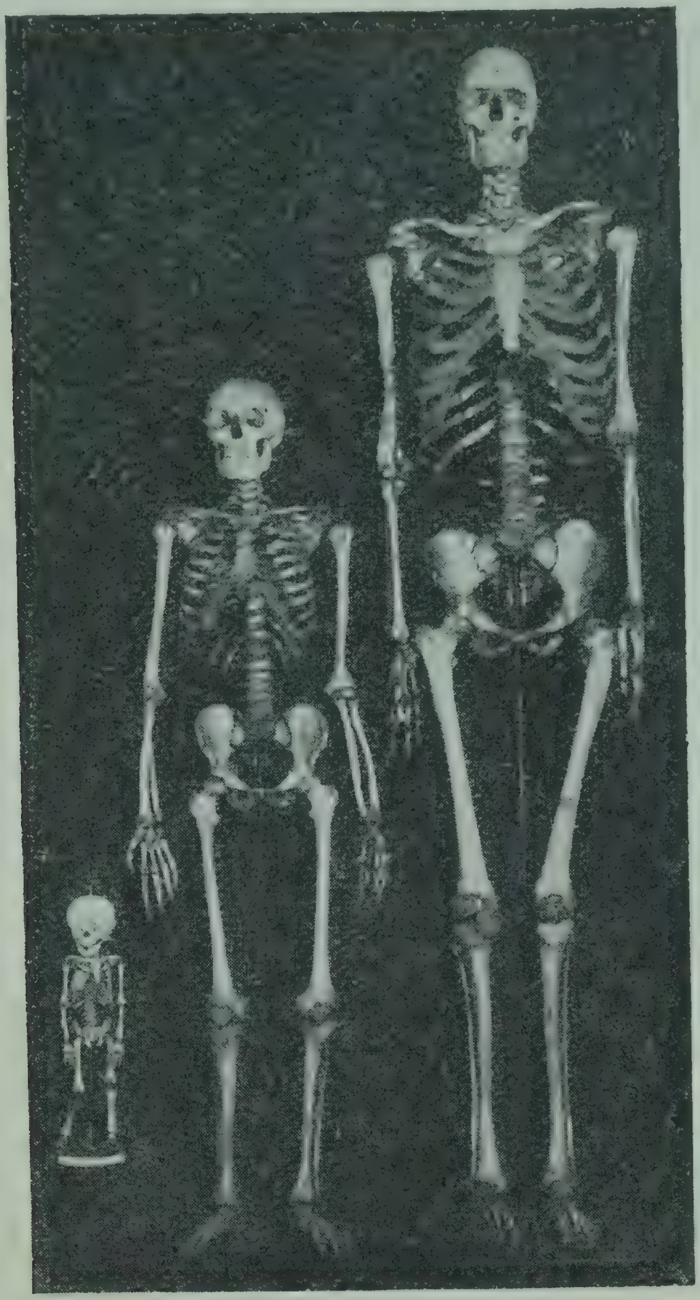


Fig. 652. Skeletons of Dwarf, Normal Man and Giant. *Dwarf*, Caroline Crachami (Cat. No. 4175·1), height 50·3 cm. (19·8 in.), pituitary fossa width 8 mm. *Normal Man* (Cat. No. 33·1), height 169·8 cm. (5 ft. 6·9 in.), pituitary fossa width 15 mm. *Giant*, O'Brien ♂ (Cat. No. 3865·1), height 231·1 cm. (7 ft. 7 in.), pituitary fossa width 22 mm. (By courtesy of Dr. H. J. Beattie, Royal College of Surgeons of England.)

Gigantism and Dwarfism. When the disease occurs during the period of active growth, the limbs and trunk bones also grow to an abnormal size resulting in a condition of "*gigantism*" (Figs. 650, 652). In such giants, X-ray examination usually shows the presence of an abnormally large pituitary fossa, and there seems little doubt that this condition is due to an over-action by which the growth hormone is produced at an abnormally high rate.

Reduced activity of the eosinophile cells is associated with the converse condition of "*dwarfism*" (Figs. 651, 652), though the condition is usually complicated by the results of deficiency in other pituitary hormones besides the growth factor. In the form known as the *Lorain-Levy* type there is a state of sexual infantilism; the dwarfs are of childish stature and proportions, though of normal intelligence; in girls, menstruation is delayed or absent. Dwarfism of the Lorain-Levy type is said to be associated with chromophobe adenomata, the depression of the hormone production being then ascribed to the pressure which results.

---

* H. M. Evans. *J. Amer. Med. Assoc.*, 1935, **104**, 1232.

† Marie. *Rev. Med.*, 1886, 298; *Brain*, **12**, 59.

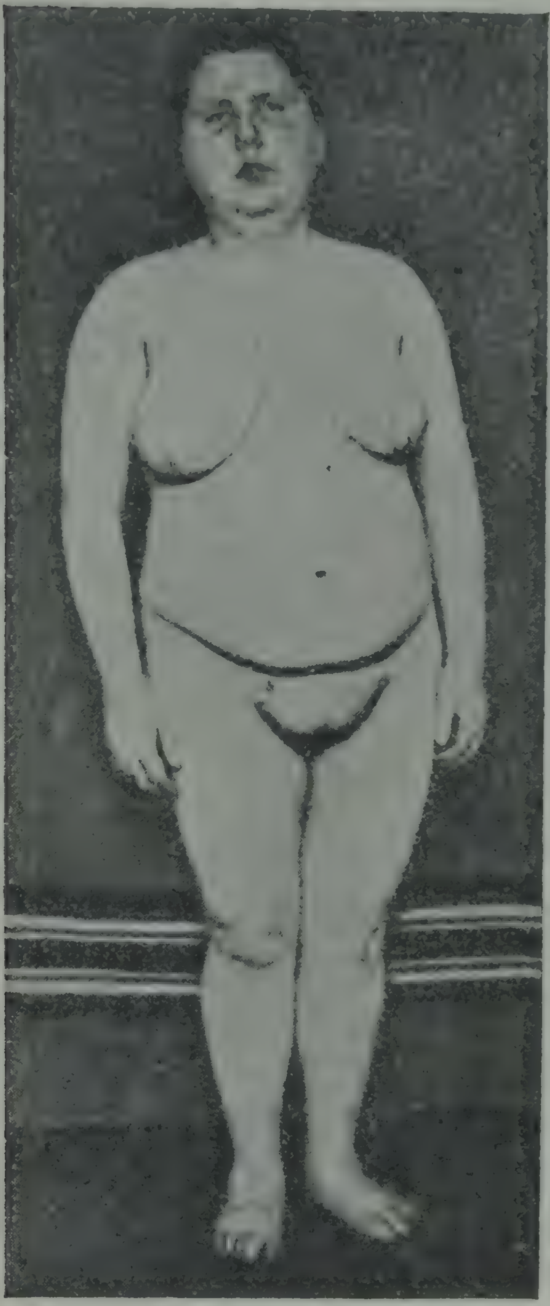


Fig. 653. Fröhlich's Syndrome. (From Deusch. *Deutsch. Zts. f. Nervenheilkunde*, 1925, **86–87**, 117.)

In the *Fröhlich-syndrome* (Fig. 653) there is adiposity and great apathy, in addition to infantilism, and sexual development is conspicuously delayed in males. When the condition arises in adult life, the conspicuous features are fatness and loss of sex functions. It is usually due to a congenital tumour of the hypophyseal duct and is of complex symptomatology.

Dwarfism may also be produced experimentally in animals by removal of the anterior lobe of the pituitary body when young. These hypophysectomised animals, of course, also show results of deprivation of other pituitary hormones, just as do some of the human cases of pituitary hypofunction. Thus some are sexually infantile (lack of gonadotropic hormone), and these retain infantile appearance and proportions. Others, in which only growth is in abeyance, have normal adult proportions. The administration of growth hormone extracts to human dwarf children has resulted in resumption of growth.*

Experimental gigantism (*e.g.* rats 50 per cent. above the normal size, and of twice the normal weight) has been produced in animals by the injection of partially purified growth hormone, prepared from extracts of the anterior lobe † (Fig. 654). The hormone can be separated in pure crystalline form.‡ Skeletal growth, checked by hypophysectomy, is restarted by the hormone, and is especially noticeable in the epiphyseal cartilages. Feeding of amphibian larvæ with fresh anterior pituitary results in accelerated metamorphosis. The growth hormone also has an effect on protein metabolism, which consists of an increased laying down of protein in muscles and a diminution of protein breakdown. Appetite is improved, but even when the food given is not increased, growth is still accelerated. Muscular tone and power are augmented by the hormone. The growth hormone may be inactivated in pituitary extracts by adding small amounts of tricresol.

It now seems certain that the growth hormone is identical with the *diabetogenic hormone* mentioned on p. 971. It has been observed that glucosuria, or even diabetes, may be associated with acromegaly, while in the Fröhlich syndrome there is increased carbohydrate tolerance; it has been shown that hyperglycæmia and glycosuria may be produced after an interval of some days by the injection of extracts of the anterior lobe. When the pituitary is extirpated, on the other hand, the animals become very sensitive to insulin and even tend to develop hypoglycæmic symptoms if not liberally fed.

---

* Taylor. *Endocrinol.*, 1938, **22**, 707.

† H. M. Evans and Simpson. *Amer. J. Physiol.*, 1931, **98**, 511; Bülbring. *Q. J. Pharm., Pharmacol.*, 1938, **11**, 26.

‡ Li *et al.* *J. Biol. Chem.*, 1945, **159**, 353.

Whether the hormone administration results in growth or in diabetes is robably dependent upon the age of the animal and on the extent to which the slet tissue can respond by producing more insulin : if it can, then growth occurs, f not, diabetes ; in young individuals the former effect is the more likely.*

It seems certain that some cases of spontaneous hypoglycæmia are due to hypofunction of the anterior pituitary,† and on the other hand, it is possible that certain clinical cases of diabetes may be due to hypertrophy of the anterior lobe ; the mode of action of the principle is uncertain (*v.* p. 974).

Although distinct from the thyrotropic hormone, the growth hormone is potentiated by the latter ; both lower the blood amino-acids, while thyrotropin lowers blood urea. The growth hormone reduces the activity of liver arginase, but the relation of this to the diminished protein metabolism is uncertain, because hypophysectomy also reduces liver arginase. Probably *in vivo* there is collaboration between growth, thyrotropic and adrenocorticotropic hormones in stimulation of growth. As explained above, the presence of an adequate supply of insulin is essential, as otherwise the growth hormone has a diabetogenic action.

(2) **Thyrotropic Hormone.**‡ Extirpation of the pituitary body leads to an involution of the thyroid, accompanied by reduction in the height of its epithelial cells, reduction of about 30 per cent. in the basal metabolic rate, and a reduced nitrogen excretion. The result is due to the removal of a hormone, the thyrotropic hormone, which exerts a stimulating effect on the thyroid. This effect probably consists in part in the facilitation of the release of thyroxine from the thyroid into the blood, since after hypophysectomy the thyroid, though smaller, contains more thyroxine than normal. Conversely, thyroidectomy results in enlargement of the pituitary body. It has been suggested that myxœdema and cretinism may be primarily due to pituitary hypofunction, the thyroid being affected secondarily. The hormone can be prepared reasonably free from active impurities,§ and when injected leads in a few days to an increase up to + 35 p.c. in basal metabolic rate, accompanied by evidence of increased formation of thyroxine and by thyroid hyperplasia‖ with increased height of its epithelial cells ; it also produces exophthalmos.¶ Assay of the hormone is made by estimating the effect of five daily doses on the weights of the thyroids of immature female guinea-pigs (200 g.) ; the weight increases from 30 mg. to anything up to 100 mg., according to the dose.** The maximal effect on metabolism is reached after one or two weeks of daily injections, and thereafter dies down, and returns to the normal, or even falls below it. The reduction appears to be due to the presence, in the serum of the injected animals, of an anti-body, due to the antigenic action of the protein, usually from a different species, of which the hormone consists, or with which it is associated.

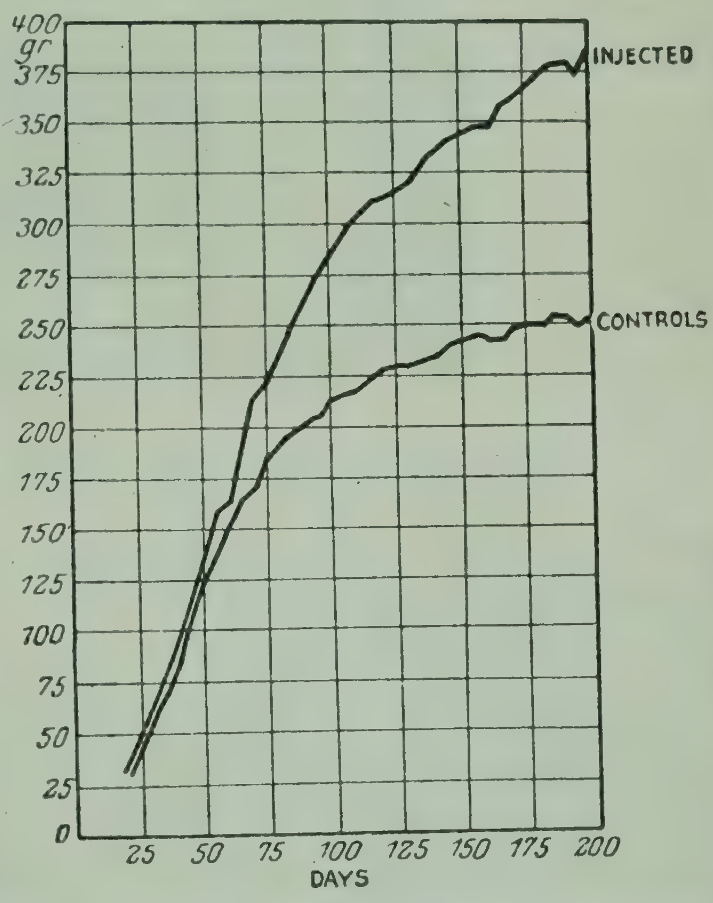


Fig. 654. Effect of Daily Injections of Anterior Pituitary Extracts on the average weights of 38 rats, as compared with 38 untreated controls. (After Evans and Long.)

* Young. *Lancet*, 1948, ii, 955.

† Best. *Lancet*, 1934, 1155, 1216.

‡ Collip. *J. Amer. Med. Assoc.*, 1935, **104**, 916.

§ Anderson and Collip. *J. Physiol.*, 1934, **82**, 11 ; Creep. *Amer. J. Physiol.*, 1935, **110**, 692.

‖ M. Krogh, Lindberg and Okkels. *Act. Path. ct. Microbiol. Scand.*, 1932, **9**, 1.

¶ Pochin. *Clin. Sci.*, 1944, **5**, 75.

** Rowlands and Parkes. *Bioch. J.*, 1934, **28**, 1829.

The hormone produces no effects on thyroidectomised animals. Thyrotropic hormone is excreted in the urine, and may be recognised there after thyroidectomy; this may be due to increased production or to failure of fixation by the thyroid.

(3) **Adrenocorticotropic Hormones.**[*] After hypophysectomy there is considerable atrophy of the inner layers of the suprarenal cortex, which can be reversed by injection of anterior pituitary extracts, and which has its counterpart in the clinic in the small suprarenals seen in pituitary dwarfism. The loss of muscle tone following hypophysectomy is probably due to adrenal cortical atrophy, since it is restored by administrations of cortex extracts. A reciprocal connection is also indicated by the fact that after adrenalectomy and in Addison's disease, there is a reduction in the number of basophile cells in the anterior pituitary.

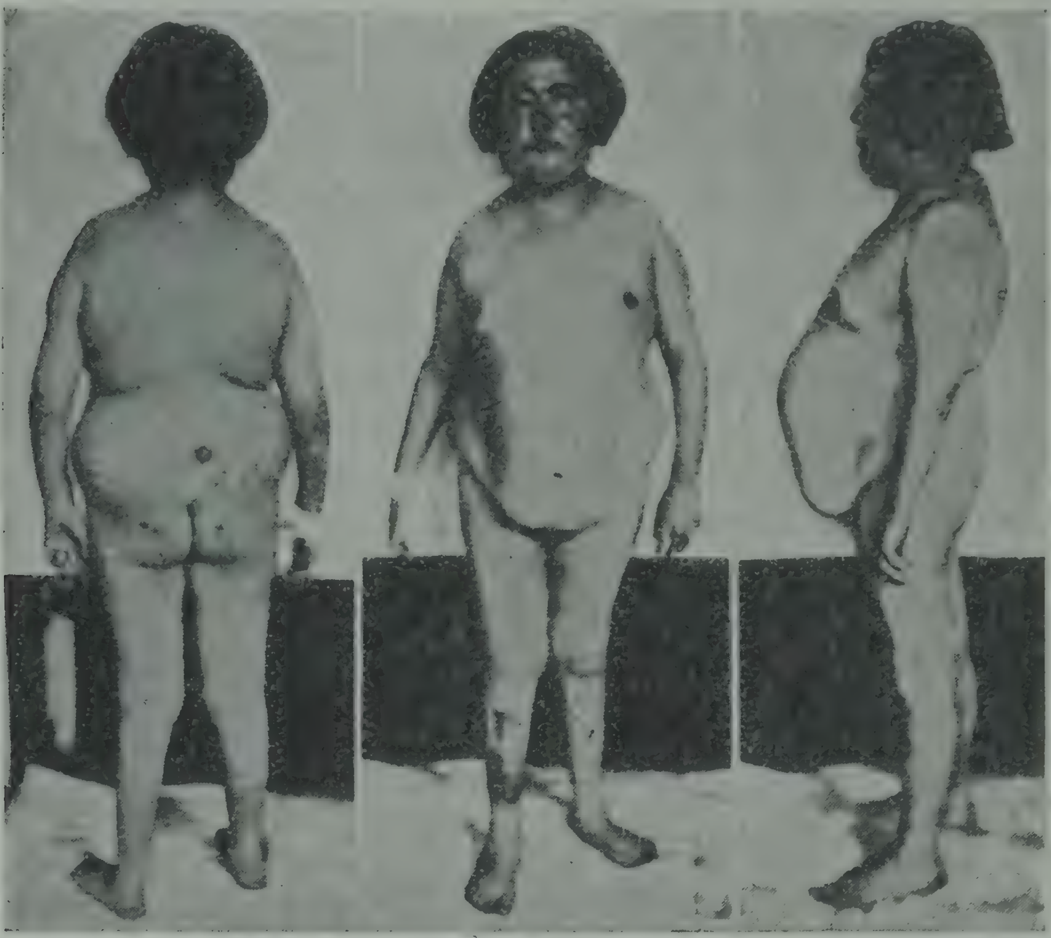


FIG. 655. Case of pituitary basophilism. Note the dark-coloured face and adiposity. (From CUSHING, *Bull. Johns Hopk. Hosp.*, 1931, **49–50,** 137.)

A pure protein, free from other pituitary hormones, has been prepared.[†] The extracts improve the resistance of rats to exposure to cold, but this effect is not obtained if the rats are previously adrenalectomised. Hypertrophy of the suprarenal cortex can be obtained by the injection of such extracts, and cortical hyperplasia is not infrequently seen in acromegaly. Cushing showed that in certain cortical adrenal tumours, associated with painful adiposity and reproductive disturbances, with hirsuties in females, there is frequently a basophile adenoma of the pituitary [‡] (Fig. 655).

Since it has been shown that the suprarenal cortex has important effects on the metabolism of carbohydrate, fat and protein, caution must always be exercised in concluding that metabolic effects of pituitary extracts are directly attributable to the action of pituitary metabolic hormones (*see* below) and not to indirect effects of the adrenocorticotropic hormone. An effect on protein metabolism is probably of this nature. Hypophysectomy lowers liver arginase, while corticosterone increases it in normal, hypophysectomised or adrenalectomised rats; the adrenocorticotropic hormone also increases liver arginase.[§]

---

[*] COLLIP. *J. Amer. Med. Assoc.*, 1935, **104,** 920; COLLIP *et al.* *Lancet*, 1933 (ii.), 347.

[†] SAYERS *et al.* *J. Biol. Chem.*, 1943, **149,** 425.

[‡] CUSHING. *Lancet*, 1930, pp. 119, 175.

[§] FRAENKEL-CONRAT, SIMPSON and EVANS. *Amer. J. Physiol.*, 1943, **138,** 439.

(4) and (5) **The Gonadotropic Hormones.** It was discovered by Smith and Engle, and by Zondek and Aschheim in 1927, that the subcutaneous mplantation of fresh anterior pituitary gland into young animals led to precocious sexual development, as a result of which the gonads not only matured early, but also, in consequence of their maturation, produced other hormones having important effects on the secondary sex characters and on the sex cycle. These hormones will be referred to in more detail in Chapter LII.

(6) **The Lactogenic Hormone.**[*] Various other factors, to be dealt with later (Ch. LIII), are responsible for the initial growth and development of the mammary glands during pregnancy, but these effects do not culminate in secretion of milk. Once the growth of the glands has reached a suitable stage, however, secretion can be initiated and maintained by a hormone formed in the anterior pituitary. The name "*prolactin*" has been given to this hormone, which can be separated from the other pituitary constituents, and shown to be independent of them. Pure prolactin is a crystalline protein,[†] rich in sulphur, and of potency of 30 I.U. per mg.[‡]

It is assayed by its effect in causing a growth in the crop-gland of pigeons. The international unit of prolactin has been defined as the specific activity contained in 0·1 mg. of a standard preparation.

FAT METABOLISM. The presence of a hormone affecting fat metabolism was first indicated by the work of Burn and Ling,[§] who injected NaOH extracts of anterior pituitary into rats fed on butter. The injections caused considerable increase in the excretion of acetone bodies. It has since been shown that there is also an increase in the ketone bodies, particularly the β-hydroxybutyric acid, of the blood. The name "*ketogenic hormone*" has accordingly been given to the responsible principle. One of its effects seems to be an increase in glycogen storage in the liver; another is the production of fatty livers in fasting guinea-pigs—an effect which lipocaic can prevent.[||] It acts on thyroidectomised animals and appears to be distinct from the other pituitary hormones. It may be identical with a factor causing increased metabolic rate and fall of R. Q.[¶] It is probable that the close resemblance to clinical diabetes which results from dosage with the anterior pituitary is usually due to the combined actions of the carbohydrate and ketogenic hormones.

THE POSTERIOR LOBE OR NEUROHYPOPHYSIS

Extracts made from the neural lobe and neural stalk of the pituitary cause a rise of blood pressure when injected intravenously,[**] and this observation formed the starting point for investigations of the active principles of the posterior lobe, with its attached parts.[††]

The physiological effects which can be produced by injection of pituitary "posterior lobe" extracts ("pituitrin") containing all the hormones (and also histamine as an impurity) are as follows:

(1) Rise of arterial blood pressure (*pressor* action) with cardiac slowing, when large doses are injected intravenously into the anæsthetised animal. A second dose soon after the first produces a fall of pressure. Most arterioles are constricted. Injected subcutaneously (0·5 c.c.) in man, there is no rise of blood pressure, but intense pallor due to cutaneous vaso-constriction. The cardiac output may be diminished and the coronary vessels are constricted. According to Krogh, a certain content of pituitrin in the circulating blood is essential for the maintenance of the normal tone of the capillaries in the frog.

[*] RIDDLE. *J. Amer. Med. Assoc.*, 1935, **104**, 636.
[†] WHITE, CATCHPOLE and LONG. *Science*, 1937, **86**, 82.
[‡] LI *et al.* *J. Biol. Chem.*, 1942, **146**, 627.
[§] BURN and LING. *J. Physiol.*, 1928, **65**, 191; 1930, **69**, 19P.
[||] JULIAN *et al.* *Amer. J. Physiol.*, 1943, **138**, 264.
[¶] O'DONOVAN and COLLIP. *Endocrin.*, 1938, **23**, 718.
[**] SCHÄFER and OLIVER. *J. Physiol.*, 1895, **18**, 230.
[††] GEILING. *J. Amer. Med. Assoc.*, 1935, **104**, 738.

(2) Contraction of plain muscle generally. Administration of pituitar extract to a lactating animal leads to sudden expulsion of milk, due to contraction of the plain muscle in the walls of the milk sinuses. There is als contraction of the plain muscle of the alimentary canal. The production o contraction of plain muscle of the uterus is called the *oxytocic* action.

(3) Expansion of the melanophores (pigment cells) of the frog's skin (Hogber and Winton), so that the skin looks darker in colour. (Especially by extract of the pars intermedia.)

(4) *Diuretic Effect.* When urine of high osmotic pressure is being secreted injection of pituitary extracts often causes diuresis, with increased excretion o Cl' and reduced excretion of PO_4'''. Also, in anæsthetised animals, a brie diuresis follows the injection of large doses : this is due to the rise of bloo pressure.

(5) *Anti-Diuretic Effect.* When urine of low osmotic pressure is being produced, pituitary injections have the opposite effect on water excretion though Cl' excretion is increased. Thus, if injected into a normal man wh has received a large draught of water (1 l.) pituitrin delays for four or five hour the excretion of the water by the kidneys, so that it accumulates for a tim in the plasma and in the tissue fluid in equilibrium with it. Ultimately a pathological condition, water poisoning, may develop owing to excessive retention of water. The effect appears to be due to an acceleration of water reabsorption in the kidney tubules. It also increases the excretion of certain ions, *e.g.* Cl, as has been shown with the heart-lung-kidney preparation (p. 1013) The principle causing this water retention is much more abundant in mammalian than in other pituitaries * and is absent from the pars intermedia. The condition of diabetes insipidus, however brought about, may be relieved temporarily b the subcutaneous injection of a watery extract of the posterior lobe.

(6) Hyperglycæmia, glycosuria, lowered sugar tolerance, loss of liver glycogen, and an action antagonising that of subcutaneously injected insulin.

There is also a rise in the fat of the liver, and a decrease in basal metaboli rate. To what extent these effects are due to a specific hormone or to hormones from the anterior lobe is uncertain.

The best known actions are the pressor effect, the anti-diuretic effect, and the "oxytocic" action on uterine muscle. For all these purposes pituitrin in subcutaneous injection, is employed clinically.

It had long been disputed whether these actions are to be attributed to a single active principle, or whether more than one such hormone is concerned as was claimed by Dudley. Kamm and others † seemed to have settled the discussion in favour of the latter, and the separated oxytocic and presso principles ("*pitocin*" and "*pitressin*") can now be obtained on the market. The cardiovascular and anti-diuretic effects, and the contraction of most of the plain muscle of the body are due to the pitressin fraction. The contraction of the uterus and the diuretic effect ‡ are specific effects of the oxytocin, but the human uterus is not always equally sensitive to its action ; in the early stages of pregnancy it is ineffective, but as pregnancy advances the sensitiveness increases and reaches a maximum at parturition ; after delivery the uterus is again insensitive for some time. More recently, the idea of a single polyvalent active substance, which is a protein, has been revived.§ The smaller molecules of pitocin and pitressin which are separable from extracts are, on this view, split off by decomposition of the protein molecule.

The melanophore-expanding principle is a separate substance, derived from the pars intermedia.

The unit of pressor activity is that of 0·1 c.c. of a standard pituitary

---

* Heller. *J. Physiol.*, 1941, **99**, 246.

† Kamm *et al.* *J. Amer. Chem. Soc.*, 1928, **50**, 573.

‡ Fraser. *J. Physiol.*, 1942, **101**, 236.

§ Irving and Du Vigneaud. *Ann. New York Acad. Sci.*, 1943, **43**, 273.

(posterior lobe B.P.) extract. Milliunits are commonly employed in standardising, the tests being made on spinal cats.

The international unit of oxytocin is the activity of 0·5 mg. of standard dried posterior lobe, as tested on an excised guinea-pig's uterus under standard conditions.* The oxytocic principle has been concentrated by adsorption methods † and an activity of 225 units per milligramme reached; this gives a powerful contraction of the uterus in a concentration of 1 in 550 millions.

Functions of the Posterior Lobe. There now seems to be little doubt that the posterior lobe has as its main function the control of the water balance of the body, by regulation of the secretion of urine. This may be effected by an appropriate interplay between the anti-diuretic properties of the pitressin fraction and the diuretic effects of the oxytocin fraction. If so, we have yet to learn how these antagonising factors are severally released. As a subsidiary function, it is regarded also as playing some part in the conduct of parturition. The pressor action is probably incidental, and is only seen when relatively large amounts of the pitressin fraction are injected.

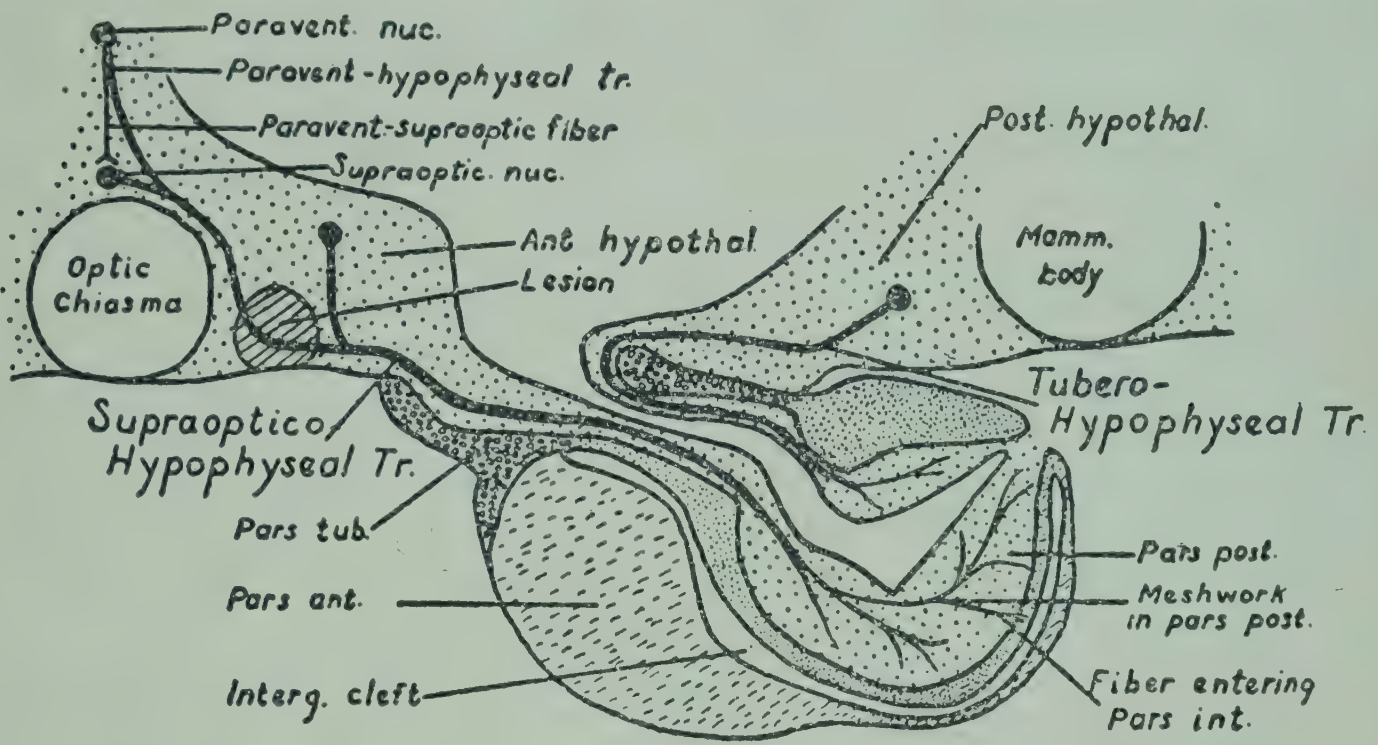


FIG. 656. Diagram of Medial Saggital Section of Hypothalamic Region of Cat showing Nervous Connections of Posterior Lobe of Pituitary Body. Bilateral Damage to the shaded area marked "lesion" causes Diabetes Insipidus. (RANSON, Harvey Lecture, 1936–37.)

DIABETES INSIPIDUS. In some disorders of the pituitary body or hypothalamus, diabetes insipidus occurs. This is a state in which large volumes, 10 to 30 l. daily, of dilute urine are passed; the urine shows no other abnormality than an excess of water; the large water output causes intense thirst and necessitates the drinking of large volumes of water. Injection of pituitary extract causes a large reduction in the urine flow, accompanied by a fall in glomerular filtration rate and in renal blood flow (*v.* p. 1014), and an increased glomerular filtration fraction.‡ In experimental animals, diabetes insipidus can be produced in one of two ways, viz. by extirpation of the pars nervosa, or by bilateral section of the supraopticohypophyseal tract (Fig. 656).

When the tract is divided, polyuria occurs at once, lasts for about five days, and returns to stay permanently from about the twelfth day. It is then found that the supraoptic nuclei have degenerated and the pars nervosa has atrophied, and contains no pituicytes, and little or no diuretic or other active principle. This suggests that the posterior lobe is dependent for its function, and indeed for its existence, on the innervation which it receives by this tract.

---

* BURN. Methods of Biological Assay, 1928. (Oxford: Med. Publ.)

† GULLAND and NEWTON. *Bioch. J.*, 1932, **26**, 337.

‡ WINER. *Arch. Int. Med.*, 1942, **70**, 61.

When the posterior lobe, *i.e.* the pars nervosa, together with the median eminence and infundibular stem is extirpated, diabetes insipidus also follows, provided that no other damage is done, and that the extirpation is complete. But *complete hypophysectomy does not produce this result*, though if an animal is completely hypophysectomised and then has anterior lobe tissue grafted under the skin, the polyuria occurs and lasts as long as the graft survives. It is therefore evident that in order to produce the state of diabetes insipidus the presence of the anterior lobe is as important as the total absence of functional pars nervosa, and it is immaterial whether the latter is made ineffective by actual removal, or by withdrawal of its nerve supply followed by atrophy. The condition is temporarily cured by the injection of small doses of pitressin.

The suppression of the diuresis by removal of the anterior lobe has its counterpart in the fact that the injection of anterior lobe extracts into normal animals often causes polyuria. This is possibly due to the general effects on metabolism, and on other endocrine organs, among which the suprarenal and thyroid must be counted.

Thyroid extracts also exert a diuretic action, and after thyroidectomy the urine formation is reduced.

WATER REGULATION. When a large volume of water is drunk there is a lag in time between the occurrence of dilution of the blood and the onset of diuresis, and it has been supposed by Verney that this is due to the time taken for the anti-diuretic principle, present to begin with in the blood, to be appropriately diminished. For there is evidence that when the water of the body is reduced, there is a secretion of anti-diuretic substance into the blood; because such substances can be recovered from the urine of thirsting animals, except when they have been previously hypophysectomised,* or when their supraopticohypophyseal tracts have been severed.†

The hormones of the pituitary posterior lobe are probably set free in response to a raised osmotic pressure of the plasma (p. 1014), but there are results which suggest that it may also be reflex. When the central end of the divided vagus nerve was stimulated, in the isolated perfused dog's head, an anti-diuretic principle could be demonstrated in the venous blood : pressor and oxytocic substance were also set free into the blood by afferent vagus stimulation. None of these substances was found in the blood from a hypophysectomised head under similar conditions.‡ The afferent arc of this reflex was traced to the supraoptic region, and the efferent arc is no doubt the supraopticohypophyseal tract.§ As a result of afferent vagal stimulation the pituicytes are found to undergo histological changes, such as a loss of granules, so that these cells may probably be regarded as the ones which secrete the hormones of the posterior lobe.||

FUNCTION IN PARTURITION. As will be shown in a later chapter, the oxytocin of the posterior lobe may play a part in determining the onset of parturition. This is thought to be due to the fact that the uterus towards the end of pregnancy becomes increasingly sensitive to the contracting effect of oxytocin, and also possibly to the fact that at the onset of labour, there is an increased liberation of oxytocin into the blood stream. There is good evidence that after hypophysectomy delivery may be difficult.

For Reference

FISHER, INGRAM and RANSON (1938). Diabetes Insipidus and Neuro-humoral Control o Water Balance. (Ann Arbor.)

FULLER ALBRIGHT (1942). Cushing's Syndrome : Its Pathological Physiology. (*Harve Lects.*, 1942–43, p. 123.)

HARRIS (1948). Neural Control of the Pituitary Gland. *Physiol. Rev.*, **28**, 139.

---

* GILMAN and GOODMAN. *J. Physiol.*, 1937, **90**, 113.

† INGRAM *et al.* *Amer. J. Physiol.*, 1939, **127**, 544.

‡ CHANG *et al.* *Chinese J. Physiol.*, 1937, **12**, 1 ; 1939, **14**, 161 ; 1938, **13**, 269.

§ HUANG. *Chinese J. Physiol.*, 1938, **13**, 367.

|| WANG. *Chinese J. Physiol.*, 1938, **13**, 405.

PICKFORD (1945). Control of the Secretion of Antidiuretic Hormone from the Pars Nervosa of the Pituitary Gland. *Physiol. Rev.*, **25**, 573.

RESEARCH PUBLICATIONS, No. XVII. (1939). The Pituitary Gland, by 46 Authors. Edited by the *Assoc. for Research in Nervous and Mental Diseases.* (London: Baillière, Tindall & Cox.)

RUSSELL (1938). Relation of the Anterior Pituitary to Carbohydrate Metabolism. *Physiol. Rev.*, **18**, 1.

SEVERINGHAUS (1937). Cellular Changes in the Anterior Hypophysis. *Physiol. Rev.*, **17**, 556.

SWANN (1940). The Pituitary-adrenocortical Relationship. *Physiol. Rev.*, **20**, 493.

WHITE (1946). Preparation and Chemistry of Anterior Pituitary Hormones. *Physiol. Rev.*, **26**, 574.

THE THYMUS

The Thymus in man grows rapidly up to the second year of life and then less rapidly; involution begins at about the fifteenth year and in later adult life only traces remain. The involution is delayed by castration and hastened by sexual activity, malnutrition and pregnancy. In certain cases of arrested development, or of general weakness in young people, the thymus has been found to be persistent. The Hassal's corpuscles, found in the medullary part, have not had any function assigned to them. Removal of the thymus appears to have no constant effect. The injection of extracts, administration by the mouth, or grafting of thymus tissue have in most instances given no regular and outstanding results.

It has been suggested that the thymus produces a hormone the effect of which is to depress neuromuscular conduction, and thymic tumours are, in fact, found in some cases of myasthenia gravis,* the underlying cause of which is a loss of neuromuscular conduction.

For Reference

DOROTHY H. ANDERSEN (1932). The Relationship between the Thymus and Reproduction. *Physiol. Rev.*, **12**, 1.

OTHER ENDOCRINE ORGANS

The internal secretions of the *pancreas* and of the *gonads* are dealt with elsewhere (Chapters XLVI and LII).

The *Pineal Gland* has so far not been proved to have any endocrine function.

CORRELATION OF ENDOCRINE ACTIONS

The functions of the individual organs of the body are integrated in two ways—by the action of the nervous system and by the action of hormones, and in the conservation of the constancy of the internal environment (homoiostasis) in its many different aspects, sometimes one and sometimes both means of correlation are brought into play. For example, the sugar content of the blood is kept approximately constant as a result of the action of hormones, such as insulin, adrenaline, the hormones of the suprarenal cortex, anterior pituitary and thyroid, and of the central nervous system acting through its effects on the liver, suprarenal medulla and pancreatic islets. It seems almost self-evident, in such a case, that each of these factors must not only be able to influence the blood sugar level, but must also in its turn be influenced by it, and /or by all the other factors that play a part: if this were not so, we should have not correlation, but metabolic chaos.

The rate at which insulin is secreted by the pancreas is not affected by altering the total amount of pancreas within reasonable limits, *e.g.* by partial excision or by the in-grafting of as many as three extra glands, but only by the amount of sugar in the blood which reaches it.

The most remarkable examples of the inter-connections between endocrine organs are found in those instances in which the anterior pituitary plays a part, and of which we have already met with some examples. For instance, the anterior pituitary produces thyrotropic hormone, which controls the growth and activity of the thyroid, while the thyroid hormone also exercises an influence on the size and activity of the anterior pituitary. But the interaction does not stop at a mutual one between two endocrine functions, but may echo widely throughout the whole endocrine system. Thus, thyroidectomy leads to enlargement, with hypofunction, of the anterior pituitary: this results in deficient formation of gonadotropic hormones and so to depression of sexual functions, atrophy of sex organs, loss of secondary sex characters, and so on. It is also probable that the anterior pituitary is subject to indirect nervous influence by means of which further correlations of function are effected.

---

* MCEACHERN. *Medicine,* 1943, **22**, 1.; SCHWEITZER. *Exp. Med. Surg.*, 1947, **5**, 168.

A little reflection will show that the permutations and combinations of actions between the different hormones must have almost limitless possibilities, and it has in fact, as well as in fancy, been possible to postulate more than one endocrine "House that Jack Built."

Various types of functional association between the endocrine organs have been found, *e.g.* :—

(1) An organ, *e.g.* the anterior pituitary, may produce a hormone which has the effect of maintaining or stimulating another endocrine organ, *e.g.* the gonads.

(2) The hormone of one organ, *e.g.* the gonads, may moderate and reduce the activity of another, *e.g.* of the gonadotropic function of the anterior pituitary.

(3) Some hormones sensitise tissues to the action of other hormones, *e.g.* thyroxine sensitises to the action of adrenaline.

(4) Two hormones, *e.g.* that of the ovary and that of the corpus luteum, may sometimes work to reinforce each other, and in other ways may be antagonistic.

In order to establish an effect of one endocrine organ, A, on another one, B, it should be shown that (1), extirpation of B produces at least as great a hypofunction as extirpation of A, and (2) the effects of extirpation of B cannot be counteracted by implantation or injection of A.

It should be noted that although hyperfunction is seen under pathological conditions in which adenomata of the endocrine glands are formed, it is not possible to produce such results of hormone excess by the grafting of additional endocrine tissue into a normal animal. This is because a hormone usually exerts a controlling action on the growth or activity of the gland which produces it. Thus, by giving large doses of thyroid, a state of hypofunction of the thyroid, with eventual atrophy of the gland may be produced. It is for similar reasons that a graft of an endocrine organ does not take well if by doing so the animal would then be provided with an excess of it, but takes best when the organ in question is deficient in amount of activity, or has been previously excised. In other words, an endocrine organ will not grow where it is not needed. When a graft is made into a normal animal it does not lead, even immediately, to liberation of excess of the hormone; thus if the thyroid and parathyroids are grafted in by blood-vessel suture into a parathyroidectomised dog, the blood calcium rises to normal in a few hours, and stays there; but if the graft is made into a normal dog there is no alteration in the blood calcium. This means that, as in other known instances, it is the level of a particular constituent in the blood which regulates the action of the endocrine organs in releasing their secretions.

Among the more important of the functional associations between endocrine organs the following may be mentioned :—

ANTERIOR PITUITARY AND GONADS. The anterior pituitary, by its gonadotropic hormone, maintains the structure and functions of the gonads, as will be explained in a later chapter, and as a result of the normal functioning of the gonads the sexual organs and secondary sex characters are also maintained in normal condition. Reciprocally, the gonads, by their hormones, restrain the activity of the gonadotropic functions of the pituitary.

Caution is needed, however, in the interpretation of some of these results. What has actually been shown is that, say, after administration of the gonad hormones, such as the œstrogens, the substance of the anterior pituitary will be found, when implanted into an immature animal, to be less potent than it usually is in provoking precocious sexual activity. This has, as stated, been interpreted as indicating that when the gonads are actively producing their hormones, these react on the anterior pituitary in such a way as to check the production of more of the gonadotropic hormone there. But histological studies of the pituitary suggest that it is, in fact, secreting *more* actively and that the absence of the gonadotropic hormone may be due, not to slackening of secretion, but to exhaustion of the secretory reserves.

A similar criticism applies to many of the statements regarding inter-relations between endocrine organs.

ANTERIOR PITUITARY AND THYROID. By its thyrotropic hormone the anterior pituitary controls the growth and activity of the thyroid, and the thyroid, reciprocally, restrains the rate of formation of the pituitary thyrotropic hormone. After thyroidectomy the anterior pituitary hypertrophies, the basophiles increasing, and the gonadotropic, thyrotropic and lactogenic hormones diminishing.

ANTERIOR PITUITARY AND SUPRARENAL. The effect of the adrenocorticotropic pituitary hormone has already been discussed. A reverse action of the suprarenal on the pituitary is doubtful, though in Addison's disease there is a diminution in the basophiles.

ANTERIOR PITUITARY AND PARATHYROID. There is some doubt of the existence of relationship between these two glands.

ANTERIOR PITUITARY AND PANCREATIC ISLETS. We have mentioned the probable relationship between the growth hormone and the islet tissue, but it is not known what *rôle* this plays in normal life and growth.

THYROID AND GONADS. Although there is no doubt that alterations in thyroid activity affect the gonads and all that depends on them, it is not quite clear to what extent the relationship is an indirect one in which the anterior pituitary is concerned. Generally speaking, thyroidectomy causes hypofunction of the gonads (but also of the anterior pituitary), a condition which is rectified by the administration of thyroid. In normal animals small doses of thyroid stimulate the gonads to increased functional activity, but large doses depress sex functions. Castration causes in some species a state of hypofunction of the thyroid, while the ovarian follicular hormone may cause hyperthyroid activity.

SUPRARENAL AND GONADS. The important effect of the suprarenal on the sex functions has already been mentioned, and is not surprising in view of the fact that cortical hormones include some of the sex hormones, and in all cases are chemically closely related to them. Tumours of the cortex, with overaction of the cortical layers produce the remarkable cases of virilism with hirsutism in women ; surgical removal causes a complete cure. In some cases the anterior pituitary is perhaps the prime offender, as in pituitary basophilism, in which, as a secondary effect, the suprarenal cortex is involved.

Nervous Control of Endocrine Organs. The endocrine organs exert their effects on other organs, including other endocrine organs, through the blood stream, and for the most part are not under the close control of the central nervous system. The ideas of hormonal and nervous control are, however, closer to one another now than in the past, for, as we have seen, the transference of nervous action is probably effected by the agency of chemical transmitters, and this may be viewed as a kind of local endocrine activity.

The suprarenal medulla is certainly under the control of the nervous system, and is reflexly excited to activity by a variety of afferent stimuli, the efferent path being by way of twigs from the splanchnic nerves. Likewise the neurohypophysis can be stimulated electrically, either directly, or *viâ* the nucleus supraopticus, to produce antidiuretic and oxytocic hormones.* The reflex discharge of antidiuretic and oxytocic substances into the blood by vagus stimulation, already mentioned, is similar proof that the posterior pituitary is under the control of the nervous system.

---

* HARRIS. *Phil. Trans.*, 1947, **232B,** 385 ; *J. Physiol.*, 1948, **107,** 412, 430.

The islet tissue of the pancreas is another instance of nervous control, for although the secretion of insulin is mainly controlled by the blood sugar level, acting directly on the islets, there is probably a second line of action, perhaps reserved for emergencies, in which the vagi act as secretory nerves to the islet tissue.

Other instances are less definite, but in some respects the anterior pituitary provides an example. In some mammals, such as the rabbit, the maturation and discharge of ova does not occur until about twelve hours after copulation has taken place. It does not occur even then if the pituitary is extirpated within half an hour of coitus,* but can be produced at any time by the injection of an extract of the anterior pituitary, or by electrical stimulation of the tuber cinereum, though not of the anterior pituitary itself. It is inferred from these facts that the influence of the nervous system is indirect, that a humoral transmitter is liberated in the tuber cinereum, transferred to the sinusoids of the anterior hypophysis *viâ* the hypophysio-portal circulation, and stimulates this to release the hormones,† among which are the F.S.H. and I.C.S.H. The formation, and perhaps also the release, of other tropic hormones is probably effected in a similar way, so that the control of the nervous system over the thyroid and adrenal cortex is indirect. Discharge of adrenocortical hormone can also be effected indirectly through adrenaline liberation mediated *viâ* the splanchnic nerve.

For Reference

CAMERON (1935). Recent Advances in Endocrinology. (London : Churchill.)

COLLIP, SELYE and THOMSON (1940). Antihormones. *Biol. Rev.*, **15**, 1.

HOUSSAY (1936). "Association fonctionnelles entre les glandes endocrine" in ROGER and BINET'S *Traité de Physiologie normal et Pathologique.* (Paris : Masson.)

PARKES (1945). The Adrenal-Gonad Relationship. *Physiol. Rev.*, **25**, 203.

ROLLESTON (1936). Endocrine Organs in Health and Disease. (Oxford : Medical Press.)

SHARPEY-SCHAFER (1924–6). The Endocrine Organs. (London : Longmans Green.)

SWANN (1940). The Pituitary-Adrenocortical Relationship. *Physiol. Rev.*, **20**, 493.

THOMPSON (1941). Antihormones. *Physiol. Rev.*, **21**, 588.

ZONDEK and SULMAN (1943). The Antigonadotropic Factor, with consideration of the Antihormone Problem. (London : Baillière, Tindall & Cox.)

---

* FEE and PARKES. *J. Physiol.*, 1929, **67**, 383.

† HARRIS. *J. Physiol.*, 1948, **107**, 418.

B. REPRODUCTION

CHAPTER LI

THE FORMATION OF GERM CELLS

In multicellular organisms, the cells which conjugate to form a new cell, capable of developing into an individual, are of two kinds. One, which has generally a certain amount of reserve material stored up in its cytoplasm, is the female element and is called the *ovum*. The other cell, which consists of little more than nuclear material, is the male element and is called the *spermatozoon*. Both kinds of cells are derived from the germinal epithelium which can sometimes be

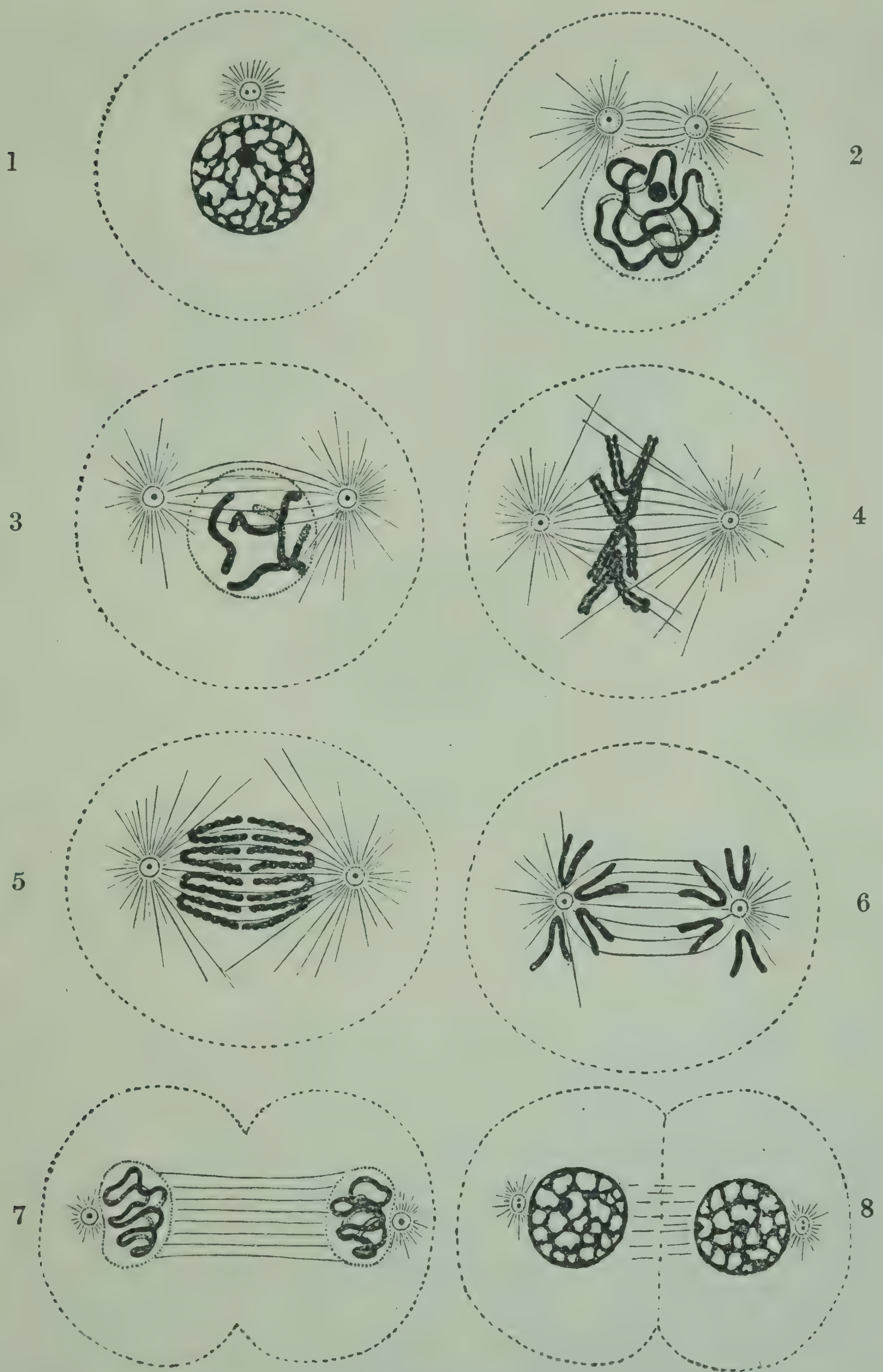


Fig. 657. Diagram showing the Changes which occur in the Centrosomes and Nucleus of a Cell in the Process of Mitotic Division. (Schäfer.)

The nucleus is supposed to have four chromosomes.

traced directly back to the first divisions of the fertilised egg. The use of the reserve material in the ovum is to serve as food for the developing individual. The ripe ovum and spermatozoon cannot be regarded as complete cells, but rather as half-cells. Before their union, or conjugation, both male and female germ cells undergo certain important changes which further differentiate them from the somatic cells of the individual. The essential differences between a ripe germ cell and a somatic cell can be best seen by a study of the nuclear changes which precede their formation.

Mitotic Division. In division, the nuclei of all somatic cells, undergo a series of changes which are similar throughout both animal and vegetable kingdoms (Fig. 657), and result in the production of qualitatively identical daughter nuclei.

The nucleus of the resting cell in its vegetative condition is generally separated from the cytoplasm by a nuclear membrane, and contains irregular masses of chromatin. In the cytoplasm of most animal cells may be seen a small particle known as the centrosome. When division is about to take place, the clumps of chromatin arrange themselves into a filament which forms a continuous skein, the "spireme stage." This then breaks up into a number of segments, often V-shaped, the chromatin filaments or *chromosomes*. Each of the filaments, in large nuclei, may often be seen to be composed of rows of granules. While this change has been occurring, the nuclear membrane in most cases disappears, and the centrosome outside the nucleus divides into two parts which travel to opposite ends of the nucleus. Round each centrosome, the cytoplasm is modified and presents a radiate appearance, the *aster*, while joining the two centrosomes is a spindle of fine fibres, the *achromatic spindle*. The V-shaped segments of chromatin arrange themselves in a circle at the equator of the spindle midway between the two centrosomes. Each of the loops then *splits longitudinally*, and each half travels towards one or other of the centrosomes, thus forming two daughter nuclei. The half-loops then join to form a skein, and may return to the condition of a resting nucleus. The entire process, as seen in tissue cultures, takes about two or three hours, though probably about twelve hours elapse between the commencement of division in a cell and the beginning of renewed division in the daughter cells. These different phases in division are presented by all somatic cells, and have received the following names:

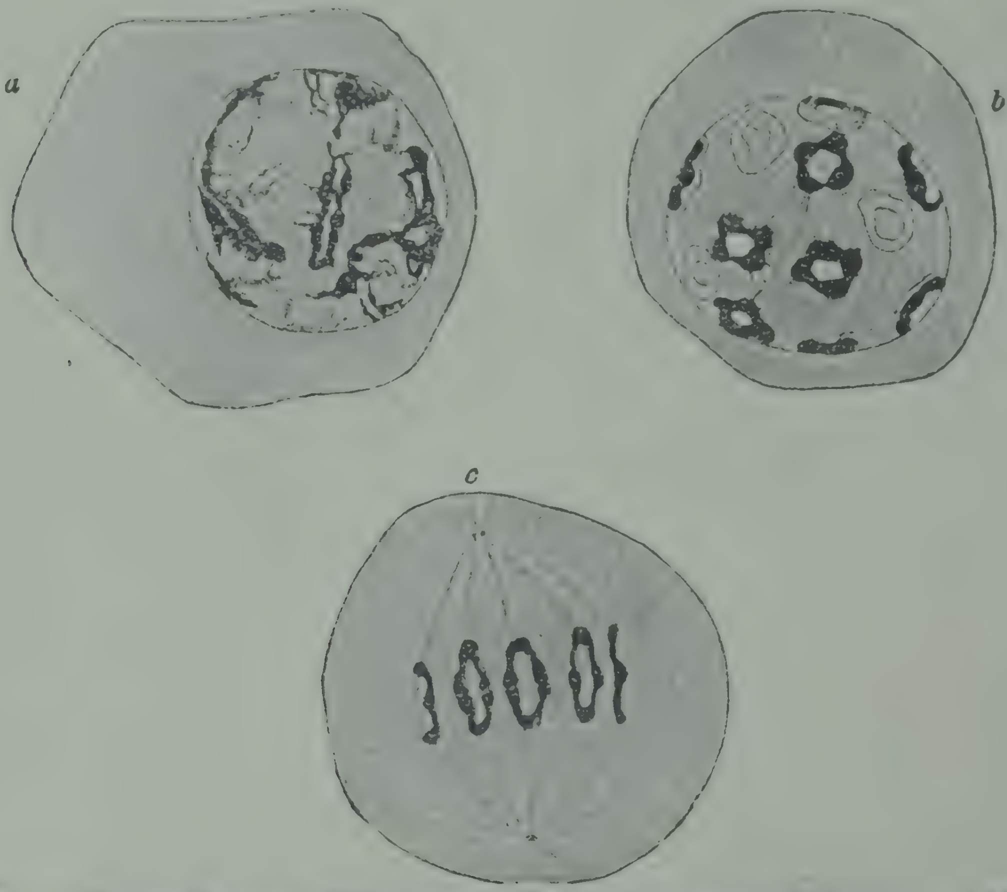


Fig. 658. Three Stages of Meiosis in Spermatocyte of Triton. (Moore.)

a. Germinal condition of chromosomes.
b. Gemini arranged in quadrate loops or tetrads.
c. Separation of tetrads into the duplex chromosomes of the daughter nuclei.

(1) *Prophase* (the formation of the spireme and of the achromatic spindle, and the breaking up of the spireme into chromatin loops or chromosomes). This phase occupies from ten to thirty minutes.

(2) *Metaphase* (the splitting of the chromosomes). This takes from two to twelve minutes.

By the action of the alkaloid colchicine the division of cells is arrested at the metaphase. This is an advantage when we wish to confirm the occurrence or increase of cell division, because after a few hours a decisive number of cells in metaphase will have accumulated.

(3) *Anaphase* (the travelling of each half-chromosome to the extremity of the spindle), which occupies about three to five minutes.

(4) *Telophase* (the retrogressive changes, leading to the conversion of the chromatin filaments into an ordinary resting nucleus, which are accompanied or preceded by a division of the cytoplasm across the equatorial part of the spindle). The latter part of the process takes about three to twelve minutes, but the retrogressive changes take from thirty minutes to two hours.

When the spireme has broken up into separate chromosomes, it is possible to count them, and it is found that the number present in any cell (*diploid number*) is constant for the species. The chromosomes are not all alike in appearance or in potentialities, but are present in pairs that are alike. Thus, every human somatic cell has 48 chromosomes, *i.e.* 24 pairs of chromosomes in its nucleus.* Identical numbers are found in types so far apart as the ox, the guinea-pig, and the onion. In the mouse, the salamander, and the lily the number is 24. Other types, such as the crustacean Artemia, are said to have as many as 168 chromosomes, the much-investigated fruit-fly, *Drosophila melanogaster*, has eight, while in Ascaris the cells contain only two to four chromosomes. There is usually an even number of chromosomes, the chromosomes of each pair being similar in size and shape. On division, each chromosome gives rise to daughter-chromosomes of exactly similar shape and size. The somatic type of cell division, called *mitosis*, seems to be adapted to ensuring an equal qualitative as well as quantitative distribution of the nuclear chromatin among the daughter cells resulting from the division of any cell. As we have seen in an earlier chapter, the nucleus, by its interaction with the cytoplasm, determines the processes of assimilation and growth of the whole cell. For the preservation

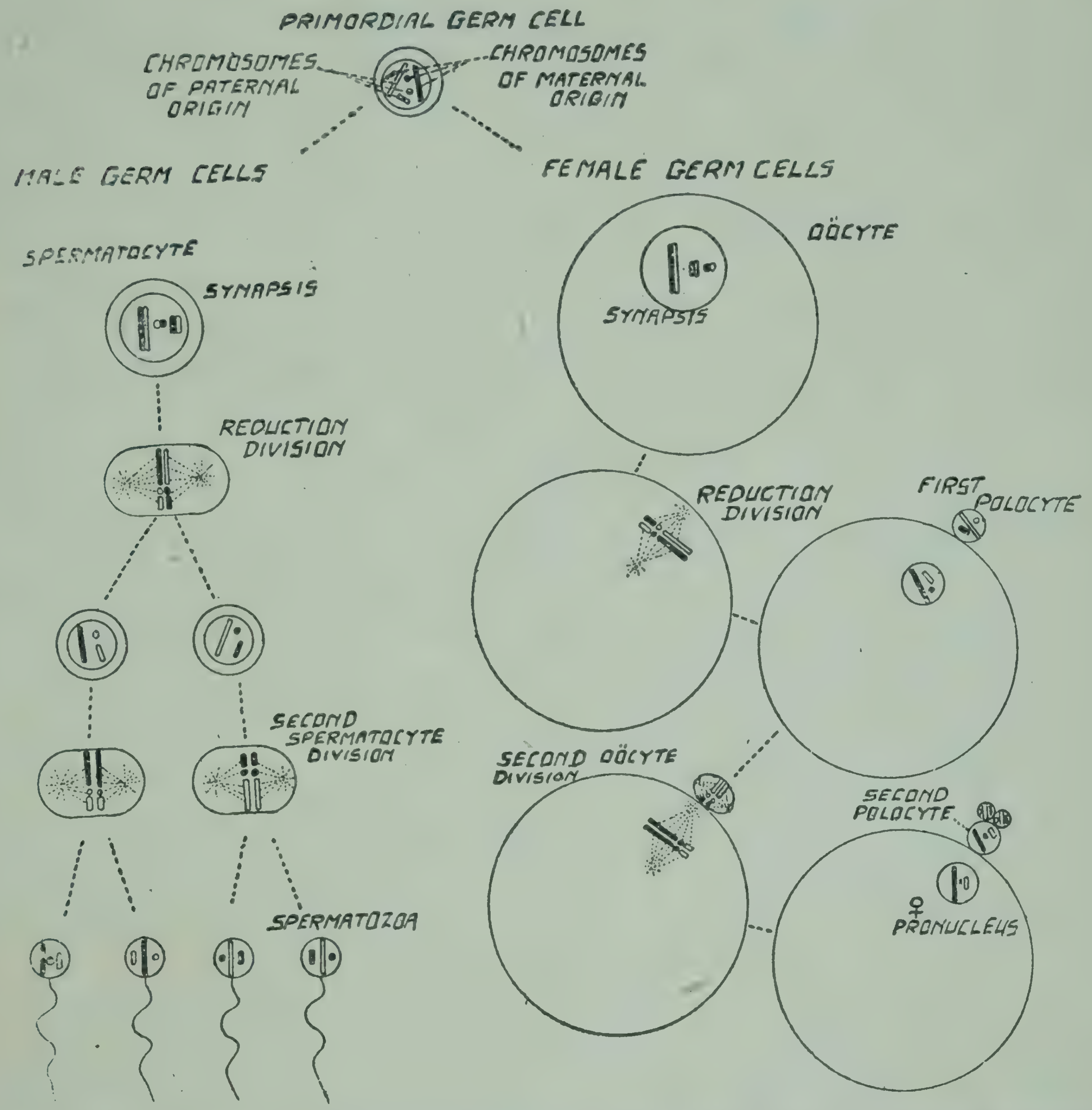

PRIMORDIAL GERM CELL
CHROMOSOMES OF PATERNAL ORIGIN
CHROMOSOMES OF MATERNAL ORIGIN
MALE GERM CELLS
FEMALE GERM CELLS
SPERMATOCYTE
SYNAPSIS
OÖCYTE
SYNAPSIS
REDUCTION DIVISION
REDUCTION DIVISION
FIRST POLOCYTE
SECOND SPERMATOCYTE DIVISION
SECOND OÖCYTE DIVISION
SECOND POLOCYTE
♀ PRONUCLEUS
SPERMATOZOA

FIG. 659. Diagram to illustrate Formation of Germ Cells with Haploid Number of Chromosomes (three). (LUDFORD).

---

* EVANS and SWEZY. *Mem. Univ. Calif.*, 1929, 9, No. 1, p. 65.

of type in cell division it is therefore essential that the nuclei of the daughter cells shall be identical in all respects with the nucleus of the mother cell.

Reduction Division. Sexual reproduction involves the conjugation of two cells, male and female, with union of their nuclei. If each of these nuclei consisted of the normal number of chromosomes, the fertilised egg cell would contain double the number characteristic of the species, and since these chromosomes would divide by splitting, the number of chromosomes in each cell would be doubled with each generation. This doubling is obviated by the fact that in the formation of the germ cells, the ovum and spermatozoon nuclei undergo a special type of division, which leads to the reduction of the chromosomes in the sexually mature cell to one-half of the number characteristic of the species. This mode of cell division is often called "reduction division," or "meiosis" (Figs. 658 and 659).

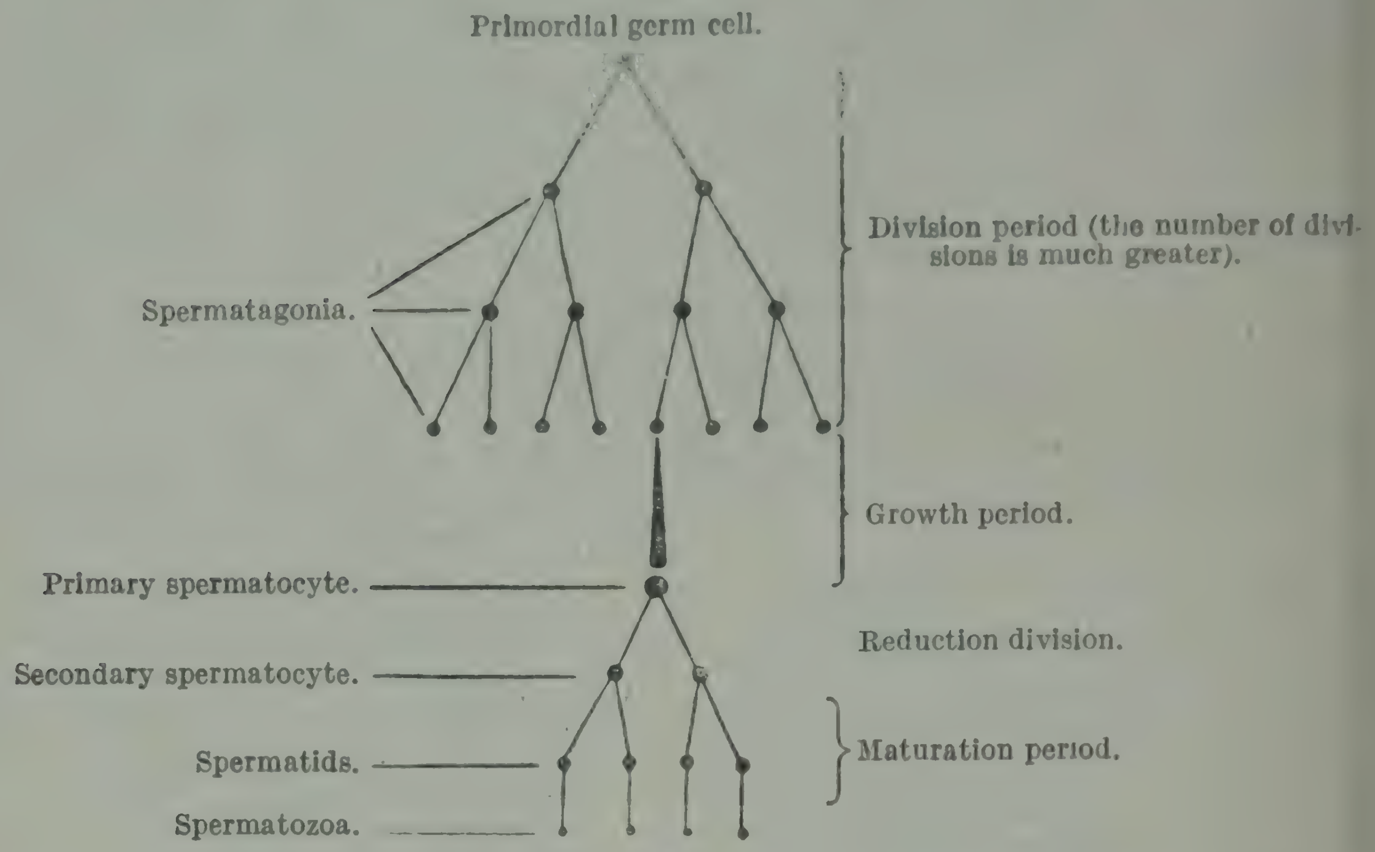

Primordial germ cell.
Division period (the number of divisions is much greater).
Spermatagonia.
Growth period.
Primary spermatocyte.
Reduction division.
Secondary spermatocyte.
Maturation period.
Spermatids.
Spermatozoa.

Fig. 660.

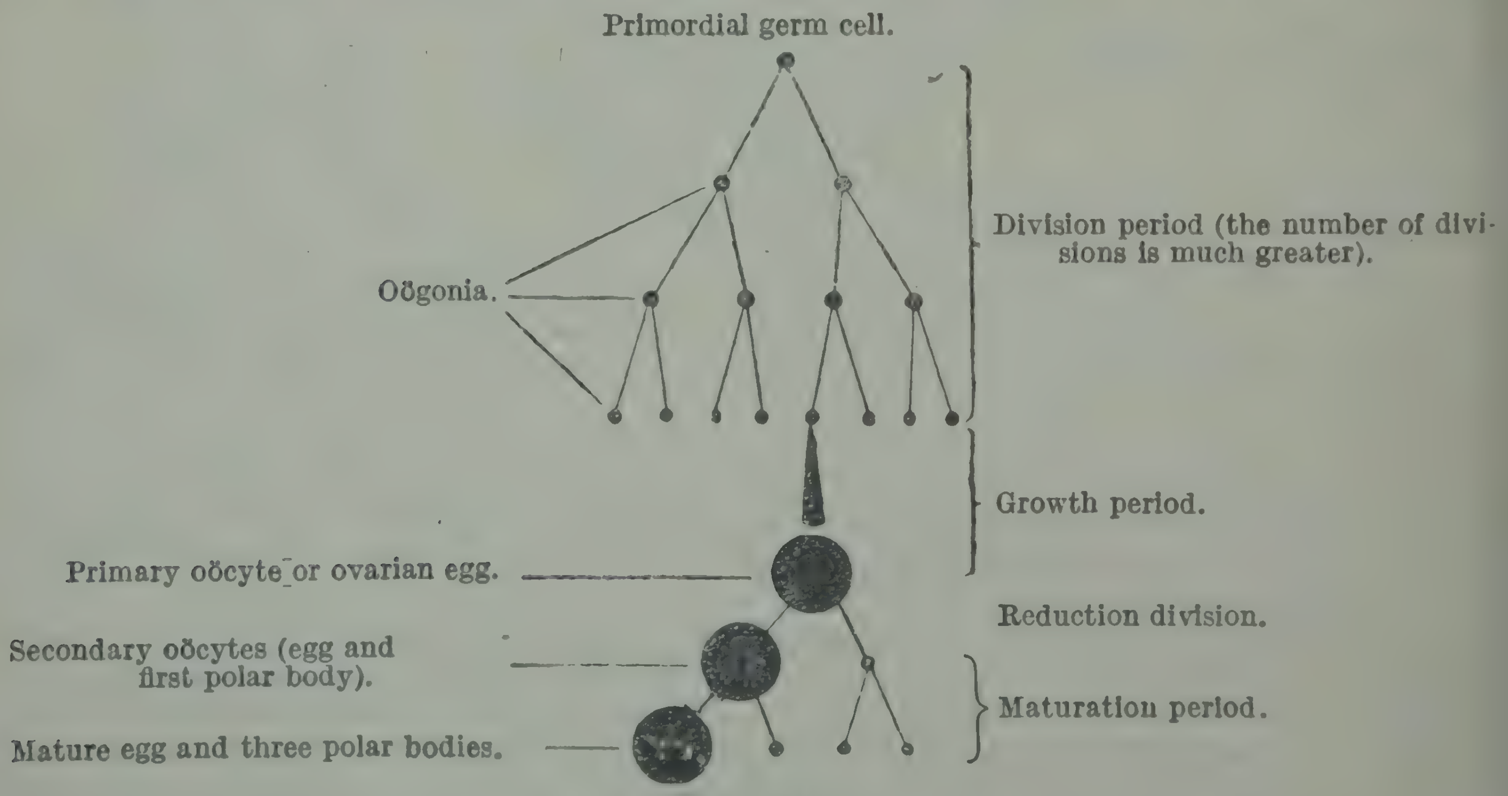

Primordial germ cell.
Division period (the number of divisions is much greater).
Oögonia.
Growth period.
Primary oöcyte or ovarian egg.
Reduction division.
Secondary oöcytes (egg and first polar body).
Maturation period.
Mature egg and three polar bodies.

Fig. 661.

We may take as an example the development of spermatozoa. The mother cells of the spermatozoa, the *primary spermatocytes*, divide twice, giving rise to four daughter cells, the *spermatids*, each of which develops into a functional spermatozoon. In the nuclear changes preparatory to the first division, the pairs of similar chromosomes become fused together (synapsis). At the first, or reduction division, the members of each pair, instead of being split longitudinally, are merely separated again, so that each of the resulting cells receives one member of each pair and hence contains what is called the *haploid* number which is only half the somatic (or diploid) number of chromosomes. At the second division, these are split in the usual way, so that the ripe germ cells also contain the haploid number of chromosomes, *e.g.* 24 in man. The distribution of the paired chromosomes between the daughter cells at reduction provides the possibility of qualitative difference between them, since the chromosome which goes to each daughter cell represents one of the two which form a pair in cells of the somatic type. If we indicate four normal chromosomes as *a*, *b*, *c*, *d*, of which *a* and *b* form one pair and *c* and *d* form the other, in ordinary somatic division each daughter cell would also contain chromosomes which may be represented as *a*1, *b*1, *c*1, *d*1, and *a*2, *b*2, *c*2, *d*2. But on reduction one daughter cell receives *a* and *c*, while the other daughter cell receives *b* and *d* in the usual way. Exactly similar changes take place when there are larger numbers of chromosomes, one member of each pair of somatic chromosomes finding its way into the ripe spermatozoon (Figs. 658 and 659).

The mitochondria and Golgi apparatus can be traced throughout all the stages of spermatogenesis, and are approximately shared out in the cell divisions. In the formation of the spermatozoa from the spermatids, well-marked changes occur in the cytoplasmic inclusions. The tail of the spermatozoon arises in association with one of the centrioles. In the mature spermatozoon the head is derived from the nucleus of the spermatid ; the acrosome, which forms a cup of variable shape according to species, is formed in association with the Golgi apparatus, in a manner somewhat similar to that of the formation of granules in gland cells. The middle piece is of cytoplasmic origin and composed largely of mitochrondria in spiral formation.

In the ovum, during maturation, analogous changes take place. Two successive cell divisions occur as in the formation of spermatozoa, but the daughter cells are of very unequal size. In the first or reduction division, the chromosome pairs separate as in the spermatocytes, so that each of the daughter cells contains the haploid number of chromosomes, and, as in the male, each of them represents one of a pair of similar chromosomes present in the somatic cells. The larger of the two resulting cells, which retains most of the cytoplasm, is still called the ovum, while the smaller one is spoken of as the " first polar body " or " first polocyte." The ovum now divides again and throws off a second polar body, the division being of the usual homotypical variety. The first polar body also divides, so that, from the original ovum, four cells are produced, one of which retains the greater part of the cytoplasm, while the others are extruded, and degenerate (Figs. 659 and 661). The only difference, therefore, between the formation of ovum and spermatozoon is that in the former case three of the cells formed by the division of the primitive ovum are abortive, whereas in the spermatozoon all four daughter cells produced from the spermatocyte remain functional. Since the nuclei of the mature ovum and spermatozoon contain only the haploid number of chromosomes, they are sometimes spoken of as *pro-nuclei*.

In oögenesis, the mitochondria increase in numbers and are evenly distributed in the cytoplasm of the ripe cell. The Golgi apparatus similarly becomes spread out, and in some cases yolk is formed in association with it ; in the mammal, however, it merely spreads out, usually near the periphery.

For Reference

Darlington (1942). The Handling of Chromosomes. (Allen and Unwin.)

Fertilisation. When the spermatozoon enters the ovum, two results normally ensue. First the egg is *activated, i.e.* provoked to cell-division. Second, *syngamy*, or the fusion of male and female pro-nuclei occurs. The essential features of fertilisation are best studied in lower invertebrates, such as ascaris or echinoderms. In the latter, fertilisation takes place in the sea-water, into which both ova and spermatozoa are extruded. The ovum of the echinoderm consists of a naked mass of protoplasm. Of the countless hordes of spermatozoa which may be in the neighbourhood of a given ovum, only one, as a rule, enters.

As soon as the spermatozoon has entered the ovum, a tough membrane is rapidly formed round the latter, so preventing the entrance of any further spermatozoa. The head of the spermatozoon enters the egg, while the tail atrophies and disappears. The head of the spermatozoon enlarges and assumes the character of a nucleus, the dense mass of chromatin breaking up first into a thread and then into the characteristic haploid number of chromosomes (Fig. 662). The egg now contains two pro-nuclei, exactly similar in appearance, one derived from the male and the other belonging to the egg itself. The two pro-nuclei approach one another and join. In many cases, there is an apparent fusion of the substance of the two nuclei. In others, the chromatin filaments of male and female simply lie side by side, forming a complete nucleus with the diploid number of chromosomes. The fate of the remainder of the spermatozoon is uncertain: it has been suggested that the acrosome plays some essential physico-chemical part in fertilisation, but there are no indications of what this may be. In some rare cases, the mitochondria of the spermatozoon have been seen to enter the cytoplasm of the ovum and mingle with those already there. In the series of divisions which precede the formation of the female pro-nucleus, the centrosome of the ovum generally disappears, whereas in the formation of the spermatozoon, the centrosome persists and forms the middle part of the spermatozoon. In many cases, the centrosomes divide in the spermatozoon itself, so that this contains two centrosomes, when it enters the egg. These two centrosomes then become the centres of attraction spheres. They diverge,

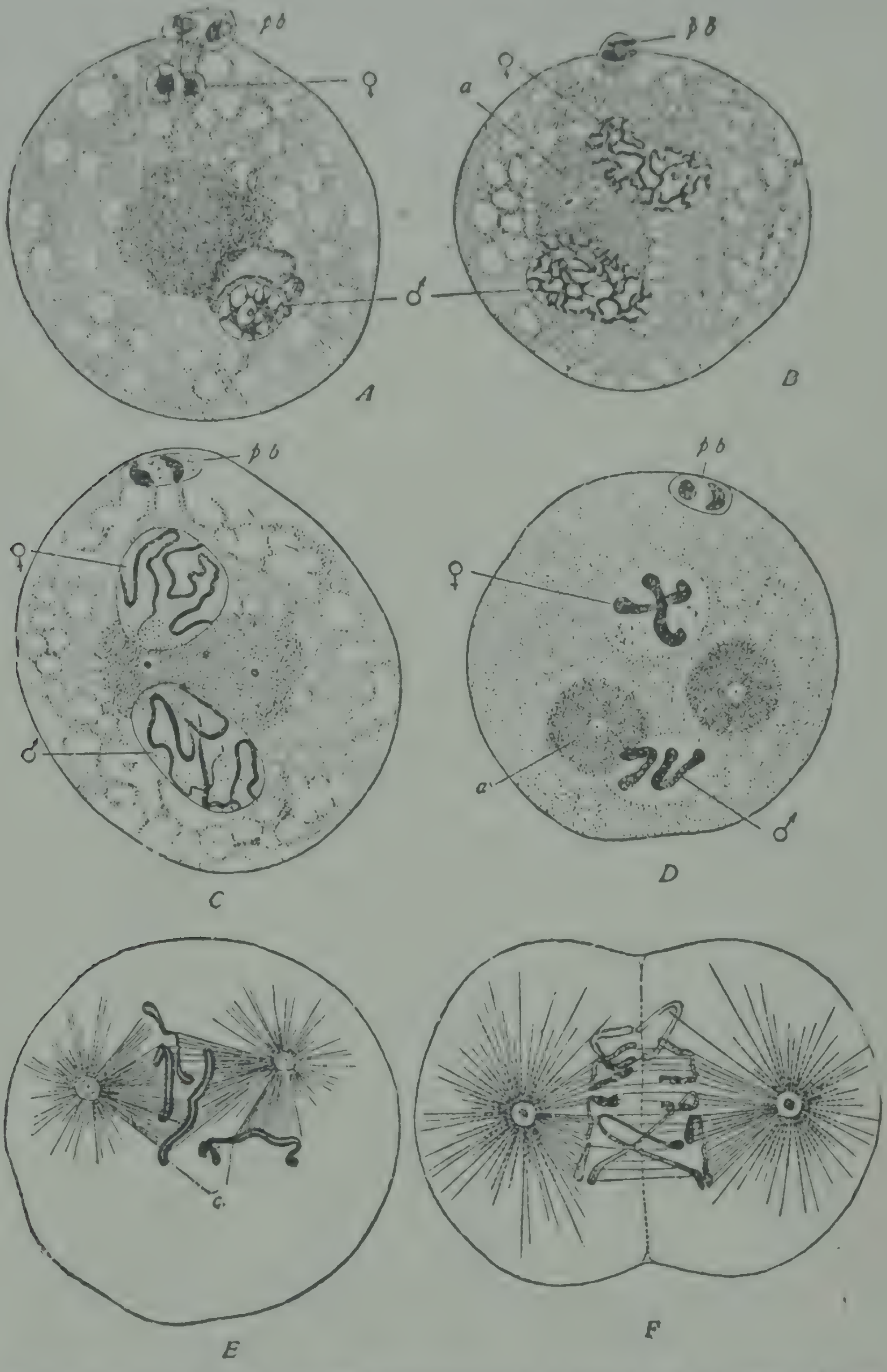


Fig. 662. Fertilisation and First Division of Ovum of Ascaris Megalocephala. (Slightly modified from Boveri and Wilson.)

A. Second polar globule just formed; the head of the spermatozoon is becoming a reticular nucleus (♂), which shows two chromosomes; just above it, its archoplasm is shown: the egg nucleus (♀) also shows two chromosomes. B. Both pro-nuclei reticular and enlarged; a double centrosome (*a*) is visible in the archoplasm. C. The chromatin is now filamentous; the centrosomes are separating. D. The chromosomes are more distinct and shortened; the nuclear membranes have disappeared; the attraction spheres are distinct. E. Mingling and splitting of the four chromosomes (*c*); the achromatic spindle is fully formed. F. Separation (towards the poles of the spindle) of the halves of the split chromosomes, and commencing division of the cytoplasm. Each of the daughter-cells now has four chromosomes; two from the ovum, two from the spermatozoon.

and between them is formed an achromatic spindle, along the equator of which the chromatin filaments of male and female pro-nuclei arrange themselves. It is doubtful, however, how far the centrosome can be regarded as a permanent cell structure.

Fertilisation is, in some species, rapidly followed by an enormous increase in the oxygen consumption of the egg, and by cell division. The egg is said to have been activated. Each of the chromosomes splits longitudinally, half going to each of the daughter cells, and this process is repeated throughout the succeeding divisions which result in the formation of the new individual. Thus, every cell of the offspring's body contains a nucleus of which exactly one half is paternal and the other maternal in origin.

Parthenogenesis. The impetus to division is normally given by the entry of the spermatozoon itself. The strong impetus to cell division given by the process of fertilisation has naturally aroused much curiosity as to its intimate character. It might be thought that, for cell division to commence, syngamy, *i.e.* the completion of the normal number of chromosomes, is essential. As against this may be adduced the fact that, in certain conditions division takes place, and an embryo is formed, without fertilisation, *i.e. parthenogenesis* occurs. The female pro-nucleus may, under certain conditions of environment or nutrition, start dividing and give rise to an embryo, each cell of which contains only the haploid number of chromosomes. In other cases of parthenogenesis, only one polar body is extruded, or the second polar body joins again with the female pro-nucleus. In these cases, the ovum contains a nucleus, with a diploid number of chromosomes, which divides and produces an individual also consisting of diploids as though resulting from the union of ovum and spermatozoon.

In unfertilised echinoderm eggs, various artificial modes of treatment will lead to the appearance of attraction spheres in the cytoplasm, and even to division of the non-fertilised egg. By the addition of mild cytolytic agents to the medium in which the eggs are contained, Loeb succeeded in imitating the changes which normally need the entrance of a spermatozoon for their occurrence. On immersing the eggs in a weak solution of lactic acid, a membrane is formed. The eggs are then placed in concentrated sea-water for a short time, and then removed to ordinary sea-water. Division rapidly occurs, with the production of an apparently normal larva. This artificial "fertilisation" is called *artificial parthenogenesis.** Loeb suggested that the activation by the spermatozoon is essentially due to a destructive action, which may be chemical in nature, on the cortical layer of the ovum, and that the spermatozoon brings with it enzymes or other chemical substances, which excite the egg nucleus and cytoplasm in the same way as the chemical measures adopted for this artificial induction of segmentation.

Mechanical damage to the outer part of the ovum may also lead to parthenogenetic development, as for instance of the unfertilised frog's egg, when punctured with a needle.

The two functions of the spermatozoon in fertilisation, viz., activation of the ovum, and syngamy, can be separated, and thus shown to be distinct. If frog spermatozoa are treated with very large doses of X-rays and then added to unfertilised eggs, these are penetrated and activated, but syngamy fails : the result is the production of morphologically uniform haploid embryos, which are in fact parthenogenetic.†

In mammals, parthenogenesis is less easy to demonstrate, but rabbit spermatozoa treated with ultra-violet light were able to activate eggs without syngamy, and cleavages resulted.‡ Rabbit ova can be activated by cooling, or by treatment with the spermatozoa of other species : segmentation occurs,

---

* Loeb. *Amer. J. Physiol.*, 1900, **3**, 135 ; *Univ. Calif. Publ.*, 1905, **2**, 83, 89, 113.

† Rugh. *Proc. Amer. Phil. Soc.*, 1939, **81**, 447.

‡ Pincus and Enzmann. *J. Exp. Zool.*, 1936, **73**, 195.

and the segmented egg can be transplanted into an animal at an appropriate stage of the œstrous cycle : in this way parthenogenetic young, all female, have been produced.*

For Reference

TYLER (1948). Fertilisation and Immunity. *Physiol. Rev.*, **28**, 180.
ROTHSCHILD, LORD (1947). *Brit. Med. J.*, ii, 239.

---

* PINCUS. *J. Exp. Zool.*, 1939, **82**, 85.

CHAPTER LII

REPRODUCTION IN MAMMALS

Hormonal Influences on the Reproductive Organs. Even within the mammalia the phenomena of reproduction show the greatest variations from species to species. Some species breed only at certain seasons of the year, others at any season. There are innumerable anatomical distinctions peculiar to species. The number of offspring produced also shows a species variation, and so do many details of the reproductive process. In the action of the hormones concerned, however, there is a good deal of common ground, and many of the observed differences between species may be traced to different quantitative or different time relations in the production of the essential hormones.

In the development and functioning of the reproductive apparatus the principal part is played by the hormones produced in the pituitary anterior lobe and in the reproductive organs themselves.

Puberty * is the transition stage on the way to full sexual maturity. The onset of puberty is due to the action of the gonadotropic hormones which now begin to be formed in the anterior lobe of the pituitary body, and which provoke the changes in the gonads which culminate in maturation of the germ cells, *i.e.* ripe spermatozoa and ova produced in the testis and ovary respectively. A further example of endocrine influence is found in the effects exerted by the hormones of the gonads themselves upon the other parts of the reproductive apparatus and upon the body generally. Thus, castration (orchidectomy). *i.e.* removal of the testes, or spaying (oöphorectomy or ovariectomy), *i.e*, removal of the ovaries, if carried out before the time of puberty, prevents the development of the secondary sexual characters, which normally occurs at this epoch.

In the human species, this period is marked, or preceded, in the male by increased growth of the skeleton, by growth of the larynx, leading to a lowering in pitch of the voice, by the growth of hair on the face and pubes, by development of the external genitalia, and by the emergence of sexual desire. The chief sign of puberty (the menarche) in the female consists in the periodic changes in the uterus, which in the human subject and the apes give rise to menstruation, *i.e.* a flow of blood and mucus from the genital organs, lasting three to five days and repeated about every four weeks. There is also enlargement of the breasts, attended by some growth of the mammary glands and by a moulding of the whole form, making it more fit for the bearing of children. The various secondary sex characters are largely the consequences of the action on the body of divers hormones which now begin to be produced in the maturing gonads. Menstruation persists, except during pregnancy and early lactation, so long as the ovary is functional. The activity of the ovary comes to an end between the forty-fifth and fiftieth year (" the climacteric," " menopause," or " change of life "). With the cessation of its activity, menstruation also stops, and the uterus undergoes atrophy. Castration after puberty, though not causing any change in the skeleton, which has already assumed its permanent form, brings about retrogressive changes in the other genital organs, analogous to those occurring in the female at the climacteric.

In animals, the series of changes known as the *œstrous cycle*, one phase of which includes the phenomena of " coming on heat," is in certain respects analogous with the menstrual cycle in the human female, and, like this, depends on the normal activity of the ovary. It is permanently abolished by extirpation of the ovaries, but may be reinduced by implantation in the peritoneum of an ovary from another animal of the same species or by injection of the appropriate ovarian hormone. This fact shows that the changes in the uterus responsible for "heat," as well as for menstruation, are independent of any nervous connections between the ovaries and the rest of the body, and must therefore be brought about by the circulation in the blood of hormones produced in the ovaries. According to some authors, the essential structures for the production of these genital hormones are the "interstitial cells" found in both the testes and ovaries of various animals. These interstitial cells are not, however, universally present.

---

* Jung. *Illinois Med. J.*, 1941, **80,** No. 6.

The chemical correlations between the ovaries and the other organs concerned in reproduction are perhaps best shown in the changes which attend *pregnancy*. This is accompanied by a great development, first of the mucous membrane, and later on of the muscular wall, of the uterus. The mucous membrane thickens, and so forms a bed for the developing fertilised ovum. With this growth of the uterus, there is a corresponding growth of the other parts of the genital tract, *e.g.* the vagina. At the same time, rapid changes take place in the mammary glands. These changes may be studied experimentally in the rabbit, in which gestation lasts about thirty days. In a virgin rabbit of a year old, it is difficult, with the naked eye, to see any trace of the mammary gland, which consists entirely of ducts lined with a single layer of flattened epithelial cells. Four or five days after fertilisation, when it is still impossible with the naked eye to discover any embryos in the swollen uterine horns, on reflecting the skin from the abdomen each mammary gland appears as a circular pink area, about 3 cm. in diameter. On section, the gland consists of ducts which are in an active state of proliferation. By the ninth day, the whole abdomen is covered with a thin layer of glandular tissue; by the twenty-fifth day, this tissue is ½ cm. in thickness and consists for the greater part of secreting alveoli. At full term, the alveoli contain ready-formed milk.

This hypertrophy of the mammary glands during pregnancy occurs after complete division of all possible nervous paths between the glands and the ovaries or uterus. If two animals are united like Siamese twins (parabiosis), and so have blood in common, and one becomes pregnant, the mammary glands of both develop and form milk. Ancel and Bouin first brought forward evidence that the corpus luteum is intimately concerned with the growth of the mammary glands, and may indeed cause a certain degree of hypertrophy of these glands in the entire absence of any product of conception within the uterus. As first shown by Heape, discharge of an ovum and formation of a corpus luteum occur in the rabbit only as a result of copulation. The same effect may sometimes be produced by artificial rupture of a ripe follicle, whereupon there is a development of the mammary glands. If no impregnation has taken place (*e.g.* if the buck has been sterilised by ligature of the vas deferens), the glands develop for fourteen days and then begin to atrophy. This period corresponds to the period of active growth of the corpus luteum. The limited growth of the glands which occurs at puberty must be connected with the growth of ripe ova or with the growth of the corpus luteum.

There are also definite relationships between the activity of the sexual organs and that of certain other endocrine organs. Thus, castration at an early age leads to persistence of the thymus gland, which normally atrophies just before the sexual organs commence their functional activity. The existence of a connection between the thyroid and the ovaries has been a popular belief for 2,000 years. In many individuals, the thyroid is perceptibly enlarged at each menstrual period. On the other hand, extirpation of the thyroid before puberty, brings about, among the other signs of cretinism, failure of development of the ovaries, so that puberty is partially or completely delayed. Hypertrophy of the pituitary body in pregnancy and in menstruation, and enlargement and increase of oxyphile cells in its anterior lobe after castration, as well as the inhibited development of the external genitalia in hypo-pituitarism, and many facts derived from laboratory experiments illustrate the fundamental relation between the pituitary body and the sex glands. There are also close relations with the suprarenal cortex, as shown, for example, in the sexual precocity associated with tumours of that tissue.

For Reference

Riddle (1931). Factors in the Development of Sex and Secondary Sexual Characteristics. *Physiol. Rev.*, **11,** 63.

THE FEMALE REPRODUCTIVE ORGANS

The essential organ of reproduction in the female is the ovary, the seat of production of the ova. The accessory organs in the mammal include the oviducts or Fallopian tubes, the uterus, in which the fertilised ovum is retained during the period of pregnancy and in which it develops, the vagina, which is adapted for the reception of the penis in the act of copulation, and the external genitalia.

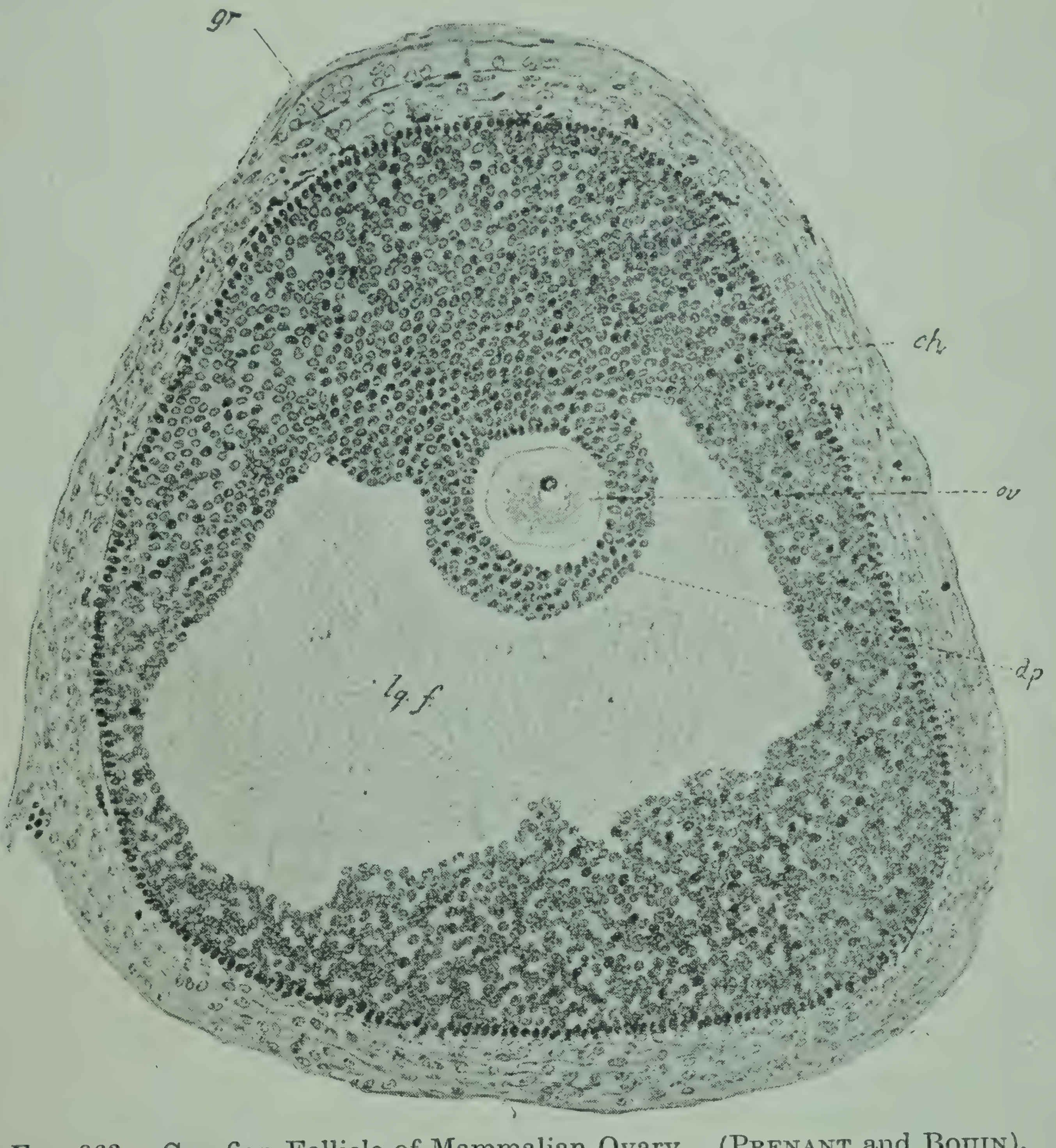


Fig. 663. Graafian Follicle of Mammalian Ovary. (Prenant and Bouin).

ov. Ovum.
ch. Theca.
dp. Discus proligerus.
gr. Membrana granulosa.
lq.f. Liquor folliculi.

Among the accessory organs we may also reckon the mammary glands, which undergo a special development during pregnancy, and serve for the nourishment of the young individual during the first period of extra-uterine life.

Ovulation. The discharge at certain periods, into the peritoneal cavity, of a ripe ovum, is called ovulation. It only occurs during the sex life. At the time of its occurrence an electrical potential difference is detectable between the cervix or vaginal epithelium and the surface of the body,* but the significance of this is not known.

---

* Burr *et al.* *Yale J. Biol. Med.*, 1937, **10,** 155.

At birth the ovary consists of a stroma of spindle-shaped cells, and is covered by a layer of cubical epithelium (the germinal epithelium) continuous with the mesothelium lining the general peritoneal cavity. Embedded in the stroma, but especially numerous just underneath the epithelium, are a vast number of "primordial follicles". These are formed, during fœtal life, by downgrowths of the germinal epithelium. They may perhaps also be formed continuously throughout the whole sexual life,* in a similar manner, though this is not the current belief. Of the cells produced in this way from the germinal epithelium, some undergo enlargement to form the primordial ova, while the others are arranged in single layers of flattened nucleated cells, the "follicular epithelium," as a sort of capsule to each ovum. Of the primordial follicles, about 70,000 are to be found in the ovary of the newborn child, and during the woman's life some 500 of these follow their full course and become ripe ova. The remainder are destroyed. During the first twelve to fourteen years of life, although some primordial follicles are already developing into well-formed Graafian follicles, with growth of the ova, these do not attain maturity, but retrogress and form "atretic" follicles.

The maturation of the ova, with their containing Graafian follicles, first occurs with the onset of puberty. The first stage in the growth of the follicle is a proliferation of the follicular epithelium, the cells of which become cubical and are arranged in several layers round the ovum. At one point in the mass of cells surrounding the ovum, a cavity appears, filled with fluid, the *liquor folliculi*. The epithelium thus becomes separated into two parts, viz. the *membrana granulosa*, several layers thick, lining the whole follicle, and the *discus proligerus*, a mass of cells attached to one side of the follicle, in which is embedded the ovum (Fig. 663). Round the growing follicle, the stroma assumes a concentric arrangement and forms a capsule, of which the internal layer consists chiefly of spindle-shaped cells richly supplied with blood vessels—the *theca interna*—while the outer layer—the *theca externa*—is made up of a tough fibrous tissue. With the growth of the follicle, the ovum also becomes larger and surrounds itself with a distinct membrane, known as the *zona pellucida*. This membrane presents a fine radial striation, which is supposed to indicate the existence of canals, through which the ovum may obtain sustenance from the surrounding cells of the follicular epithelium. The nucleus also becomes larger, and forms the *germinal vesicle*, containing a well-marked nucleolus—*the germinal spot*. The mature Graafian follicle projects from the surface of the ovary as a transparent vesicle about the size of a pea. (Its diameter is about 15 mm. in the human.) In the process of growth, the ovum has increased from a diameter of 25μ to 200μ. Before the ovum can undergo fertilisation, the double division of the nucleus, or germinal vesicle, has to take place, which leads to the formation and extrusion of the three polar bodies. This process probably occurs just before or just after the discharge of the ovum from the ovary.

With increasing size of the Graafian follicle, the membrane covering it becomes progressively thinner and more vascular, and the ovum moves towards its outer surface. At certain periods, or under certain conditions, the membrane ruptures, and the liquor folliculi, with the ovum, is discharged into the peritoneal cavity, the ovum being, at first, still surrounded by an adherent mass of the cells of the discus proligerus. In some animals, this process of *ovulation* occurs at definite periods of the year. In others, such as the ferret † and rabbit,‡ the occurrence of ovulation depends upon coitus taking place during the period of sexual activity. It is probable that, in the human female, ovulation normally occurs alternately in each of the two ovaries, once every four weeks during the thirty-five years of sexual life. Menstrual cycles not accompanied by ovulation are quite frequent, however, and probably one-fourth of cases of sterility are due to the regular occurrence of such *anovulatory cycles.* A vast number of the Graafian follicles, after developing to a certain extent, undergo regressive changes, both during childhood and during adult life. The cellular elements degenerate, phagocytes wander into the follicle and attack the degenerating ovum, so that finally the follicle is replaced by connective tissue, without the formation of any corpus luteum.

---

* Evans and Swezy. *Mem. Univ. Calif.*, 1932, **9**, 119.

† Marshall. *Q.J.M.S.*, 1904, 48.

‡ Heape. *Proc. Roy. Soc.*, 1905, **76**, **B**. 260.

The ovum, when set free, is directed into the open end of the Fallopian tube by the current set up by the cilia with which the epithelium is furnished. One factor favouring entrance of the ovum into the Fallopian tube has been said to be a movement of the fimbriated ends of the tube towards the ovary. There may be other directive forces besides this, for, when the ovary on one side is removed, and the Fallopian tube on the other side closed, pregnancy, though less likely to occur, is far from infrequent. According to Corner, the passage of the ovum along the Fallopian tube occupies three days.

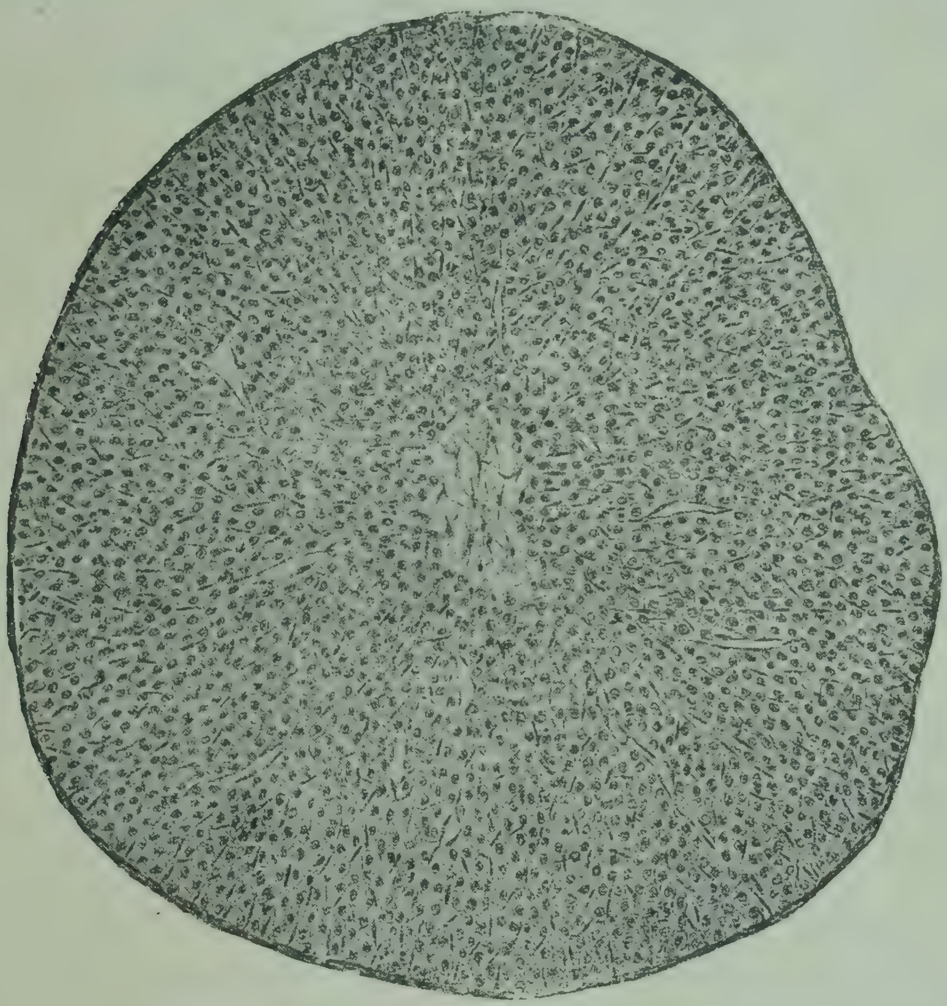


Fig. 664. Fully Developed Corpus Luteum of the Mouse. (Sobotta.)

After the discharge of the ovum, the remaining portions of the follicle undergo a characteristic series of changes, which results in the production of the *corpus luteum.* Immediately after the rupture, the cells of the membrana granulosa rapidly increase in size, a few of them undergoing mitotic division, so that a dense mass of cells is formed, nearly filling the original follicle. At the same time, the cells of the internal theca proliferate at the periphery, and the cells thus formed also grow in among the cells filling the Graafian follicle. The luteal cells finally attain a size four or five times that of the cells of the membrana granulosa in the mature follicle. Blood vessels grow from the external theca towards the centre of the follicle. The corpus luteum (Fig. 664) attains its greatest size in the non-pregnant human about the nineteenth day of the menstrual cycle, at which stage it has a grey colour: it then gradually undergoes regressive changes. The cells within the follicle then undergo fatty degeneration and present a yellow colour, due to a lipochrome known as *lutein.* If, however, the ovum, which has been discharged, undergoes fertilisation, and pregnancy results, the corpus luteum continues to grow for a considerable time and attains its largest size at about the third month, though it persists, often in an apparently degenerating condition, until after the end of pregnancy.

The big corpus luteum found in pregnancy is often spoken of as the "true" corpus luteum, and is distinguished from the *corpus luteum spurium* of menstruation, or of ovulation without fertilisation. There is no essential difference, other than that of size, between these two kinds of corpora lutea.

The Œstrous Cycle. The sexual activity of females shows recurrent changes, which vary in detail according to species, the whole cycle being known as the œstrous cycle (Heape, Marshall). In some animals, known as monœstrous, such as the bitch, there is only one œstrous period in each sexual season, of which the bitch has two a year, and between the seasons a period of quiescence, or anœstrus, is found. In others, *e.g.* mare, pig, rat and mouse, there are several heat or œstrous periods in the season, and these are separated by short resting intervals called diœstrus. In the pig, there is, at intervals of three weeks, an œstrous period lasting three days. In the human subject the alternations of rest and activity are less clearly defined. The phases of the œstrous cycle may be classified as follows :—

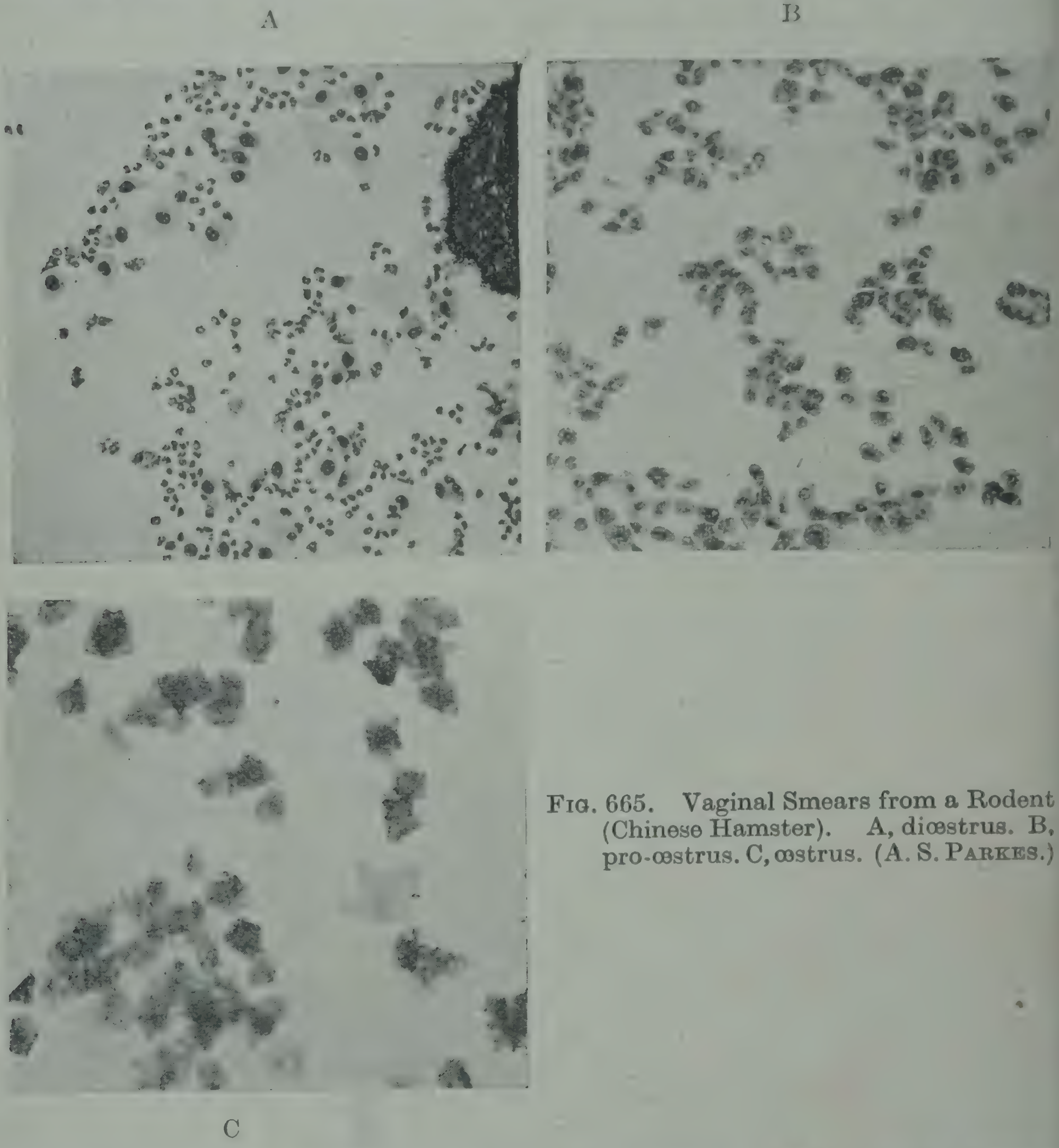


FIG. 665. Vaginal Smears from a Rodent (Chinese Hamster). A, diœstrus. B, pro-œstrus. C, œstrus. (A. S. PARKES.)

(1) Pro-œstrus, the preparatory phase.

(2) Œstrus, the period of "heat", during which alone, in most animals, copulation is permitted.

(3) Metœstrus, or luteal phase.

(4) Diœstrus (or anœstrus), or sexual rest.

These periodic modifications of sexual activity are accompanied by changes in the generative organs, in general metabolism and activity (œstrus being accompanied by increase of general activity), as well as in the behaviour towards the male, and these have been extensively studied during the last few decades.

The recurrent changes in the generative organs are normally associated with ovarian changes (follicular ripening and corpus luteum formation) ; they affect the external genitalia, the vagina, uterus, and often (especially during metœstrus) the mammary glands, and these changes are closely related to the normal sexual life in any given species. It is simplest to regard the œstrous cycle as made up of a basic *follicular phase*, consisting of a short diœstrus, a pro-œstrus and an œstrous period, with a short transition phase of met- or post-œstrus : and such a basic cycle is well seen in the unmated mouse. In most other animals, and in the mated mouse, the cycle is further modified by a *luteal phase*, due to the greater development of the corpus luteum, which, in the unmated mouse, is small and transient. As the luteal phase is connected with the functions of the corpus luteum, and may be regarded as a phase superadded to the basic follicular stage, and showing a good deal of species variability, we shall consider it later.

During pro-œstrus, growth of the follicle occurs, and at œstrus, in most animals, maturation and ovulation follow, though in some, such as the rabbit, this normally only happens if and when copulation takes place.

The study of the œstrous cycle has been greatly facilitated by the discovery, made by Stockard and Papanicolaou,* that the vaginal lining of some mammals undergoes characteristic histological changes during each phase.

In the rat and mouse, for example, in pro-œstrus the epithelial layers increase in number and in thickness, and then in œstrus the superficial cells lose their nuclei, become keratinised (cornification) and desquamate. Afterwards a transitory infiltration with leucocytes leads on to a return of the epithelium to a resting state, with few layers. Hence, by microscopical examination of vaginal smears at regular intervals, the state of the cycle at any time can be determined.

These changes, as seen in the smear, are briefly as follows (Fig. 665). In the pro-œstrous period there are nucleated epithelial cells only ; during œstrus the vagina is cornified, during metœstrus leucocytes appear in the smear, during diœstrus there are leucocytes and nucleated epithelial cells. The whole cycle in the unmated mouse takes about five days (Parkes), occupied approximately as follows : Pro-œstrus, eighteen hours ; œstrus, forty-two hours ; metœstrus, twelve hours ; diœstrus, two and a half to three days.

In the stages of pro-œstrus and œstrus the uterus becomes enlarged and congested and filled with fluid, which may be blood-stained. Its contractions become slow, large and strong, as contrasted with the feeble but frequent contractions in diœstrus or anœstrus.

The Causation of Œstrus. If both ovaries are removed the œstrous cycle ceases, either at once or after one more period. The œstrous cycle also ceases during pregnancy, and in some animals during lactation. It is not the maturation of the Graafian follicles that is the prime mover in bringing about the condition of œstrus, since Parkes has shown that in mice the œstrous cycle persists as in the normal animal after the total obliteration of the follicular system by exposure to X-rays, either before, or at any time after, birth.† It is clear, therefore, that although the presence of the ovary is essential, the presence of the ripening follicles is not. According to many investigators the essential source of the phenomena is the theca interna of follicles of all sizes. This is further borne out by the observation that in those cases where there was one last œstrous period following double oöphorectomy, the follicles of the ovaries removed had not yet reached maturity. This shows that, instead of ovulation being the cause of œstrus, they must both have some common cause, and there is clear proof that this is the gonadotropic anterior pituitary hormone.

---

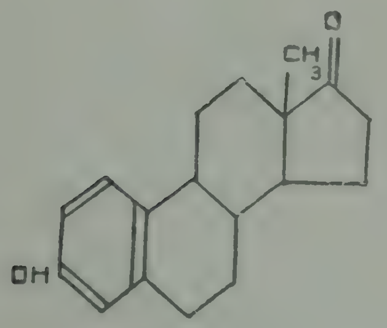
* Stockard and Papanicolaou. *Amer. J. Anat.*, 1917, **22,** 225.

† Parkes. *Proc. Roy. Soc.*, 1926, **B.100,** 172 ; 1927, **B.101,** 71, 421 ; **B.102,** 51.

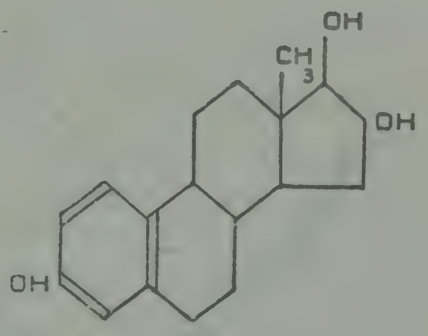
Œstrogens. It was clearly shown in 1923 by Allen and Doisy[*] that an œstrus-producing substance was present in the liquor folliculi. This, on injection into oöphorectomised mice, brought on the state of œstrus. Substances having this action are now called *œstrogens* or *œstrogenic substances.* An œstrogenic substance can be extracted from the solid parts of the ovary, from the placenta, from the urine of both men and women, but in larger amount from the urine of women at about the time of ovulation, and in still larger quantity from the urine of pregnant females. The substance is almost insoluble in water, but forms soluble alkali salts. When emulsified with water and administered subcutaneously to oöphorectomised mice, œstrus is induced in the course of the second day. Œstrogens are recognised, and extracts containing them are standardised, by the use of this method.[†] A mouse unit is the amount of a substance required to cause cornification in ten out of twenty oöphorectomised mice (but some 200 such units are needed to induce "heat"). Another test is that based on the increase in weight of the uterus of an immature mouse.

The discovery by Aschheim and Zondek (1927) that large amounts of œstrogen were present in the urine of pregnant women soon led to the isolation of one of them in a pure crystalline form. This was achieved almost simultaneously by Doisy,[‡] Butenandt,[§] and Laqueur,[‖] in 1929–30. The substance isolated proved later to be a hydroxyketone of the formula $C_{18}H_{22}O_2$.[¶] Shortly afterwards Marrian (1930)[**] isolated a second crystalline active substance which was a trihydroxy-substance. They are now called *œstrone* and *œstriol* respectively,[††] and can be separately estimated, when present together, by a colorimetric method.[‡‡] Both compounds are present in the urine of normal women (100 to 600 units daily), and of men (20 to 100 units), but are present in large amount (150,000 to 500,000 units per day) in pregnancy urine in a combined form (as glucuronides) of low œstrogenic power, from which the free and active compounds are liberated by hydrolysis with acid. Œstrone has the higher activity when tested on adult ovariectomised mice, about twice that of the trihydroxy compound. Œstriol is the more active in the induction of precocious œstrus in immature animals. A third œstrogen is present in ovarian tissue, and is called *œstradiol* ; it is the most active of them all, about six times the potency of œstrone. Which of the three compounds is the one actually elaborated by the ovary is not at present clear, though it has been supposed to be œstradiol, which is converted to the more soluble œstriol for elimination in the urine. The international units of œstrous activity are those of 0·1 μg. of œstrone or œstradiol monobenzoate (1 mouse unit = 0·1μg. of œstrone).

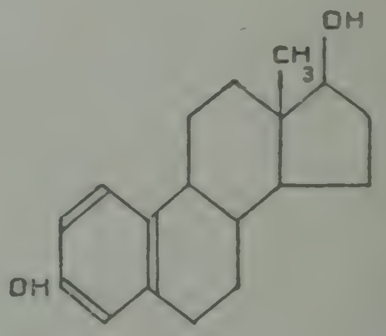
Further work indicated the following formulæ for these substances :—



ŒSTRONE



ŒSTRIOL



ŒSTRADIOL

---

[*] Allen and Doisy. *J. Am. Med. Assoc.*, 1923, **81**, 819.

[†] Marrian and Parkes. *J. Physiol.*, 1929, **67**, 389.

[‡] Doisy, Veler and Thayer. *Am. J. Physiol.*, 1929, **90**, 329.

[§] Butenandt. *Naturwiss.*, 1929, **45**, 789. *Deutsch. med. Woch.*, 1929, **55**, 2171. *Z. physiol. chem.*, 1930, **191**, 140.

[‖] Dingemanse, de Jongh, Kober and Laqueur. *Deutsch. med. Woch.*, 1930, **56**, 301.

[¶] At various times the names "*folliculin*," "*menformon*," "*theelin*" and the "*a-hormone*" have been proposed for what is probably the same hormone.

[**] Marrian. *Bioch. J.*, 1930, **24**, 1021 ; *et al.*, 1931, **26**, 25 ; 1932, **26**, 1233.

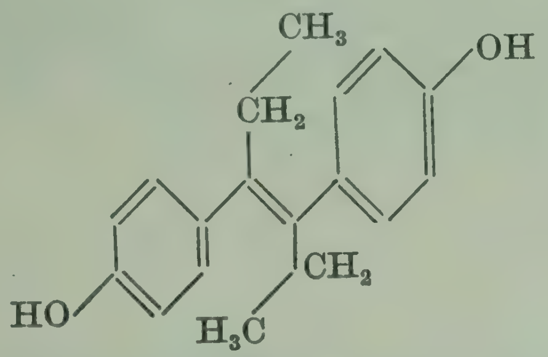
[††] Doisy has suggested the names "*theelin*" and "*theelol*" for these compounds.

[‡‡] Cohen and Marrian. *Bioch. J.*, 1934, **28**, 1603.

They are thus sterols, and closely related to some other reproductive hormones: *e.g.* progesterone and testosterone, and to adrenal cortical hormones.

They are not species-specific. Œstrone when added to the isolated guinea-pig uterus reduces the effect of oxytocin, so that smaller contractions result from it: in the pregnant uterus also it modifies the powerful contractions produced by oxytocin, retarding their rhythm and facilitating relaxation between the contractions.*

These œstrogens can be administered by subcutaneous injection in dilute alcohol, or in oil, or by implanting the solid beneath the skin. They are much less effective by mouth. The three ovarian compounds do not complete the list of naturally occurring œstrogens, for others are found in some species, *e.g.* equiline and some other sterols in the pregnant mare's urine. Also numerous synthetic substances have œstrogenic properties; some of these are sterols and some are quite unrelated, *e.g.* the most powerful is dihydroxy-diethyl stilbene (or diethyl-stilbœstrol)† which is more potent than œstradiol, though not chemically very similar.



4 : 4′-dihydroxy-α : β-diethylstilbene. "STILBŒSTROL".

The occurrence of œstrus in animals deprived of follicles by exposure to X-rays suggests that the hormone is normally produced in the stroma of the ovary. As regards its high concentration in the placenta, it is evident that, owing to the size of this organ, the placenta contains far more œstrogen than both ovaries; it has been suggested that the placenta takes up the hormones from the blood, and so protects the fœtus from their effects, but it seems more probable that they are formed in the placenta itself. Much of it is excreted in the urine of the mother.

The main effect of the injection of an adequate dose of œstrogen into the oöphorectomised female is the onset of the uterine and vaginal changes characteristic of œstrus, and also enlargement of the mammary glands. The uterus enlarges, becomes hyperæmic, and, it is stated, shows more powerful rhythmic contractions. The mammary glands (even in the male) show extension and branching of the ducts, but not of the secreting parts (Fig. 667). Large doses will check an already established lactation. Injection of œstrogen does not cause follicular maturation. The production of the hormone and ovulation, though due to a common stimulus, and normally coincident, are mutually independent.

The fate and interconversions of the œstrogens in the body are little understood.

INFLUENCE OF ANTERIOR PITUITARY. If, as indicated above, the ripening of Graafian follicles is not the cause of the presence of œstrogens, how are we to explain the periodicity of œstrus, or the occurrence at puberty of the first œstrous cycle? The work of Smith and Engle,‡ and Zondek and Aschheim § gives support to the view that the advent of puberty, as well as the rhythmicity of the œstrus phenomena, including the periodic ovulation, are part of an endocrine phenomenon, in which the anterior pituitary plays a determining *rôle*. This is effected by the gonadotropic hormones, which have since been obtained in a state of partial separation from the other pituitary hormones. Injection of suspensions of the anterior lobe of the pituitary body into adult females is followed by immediate maturation of large numbers of ova, and formation of multiple corpora lutea, or, if pregnancy follows, by the formation of abnormally large numbers of fœtuses. Injected into immature animals, the suspension brings about precocious follicle ripening and œstrus, after a latent period of about three days, *i.e.* a day longer than is required for the effect of an injection of œstrogen to manifest itself. Injection into oöphorectomised animals produces no results, which makes it evident that the effects which are produced are due to an inductive action on the ovary. In partially oöphorectomised animals, however, the hypertrophy of the remaining ovarian tissue is greatly accelerated by administration of the anterior pituitary substance. Extirpation of the pituitary body leads to atrophy of the gonads and accessory organs or, in immature animals, to failure of maturity changes.

---

* MARRIAN and NEWTON. *J. Physiol.*, 1932, **77**, 4, P.

† DODDS, LAWSON and ROBINSON. *Proc. R. S.*, 1939, **127**, **B.** 140.

‡ SMITH. *Amer. J. Physiol.*, 1927, **80**, 114; SMITH and ENGLE. *Amer. J. Anat.*, 1927, **40**, 159.

§ ZONDEK and ASCHHEIM. *Klin. Woch.*, 1927, **6**, 248.

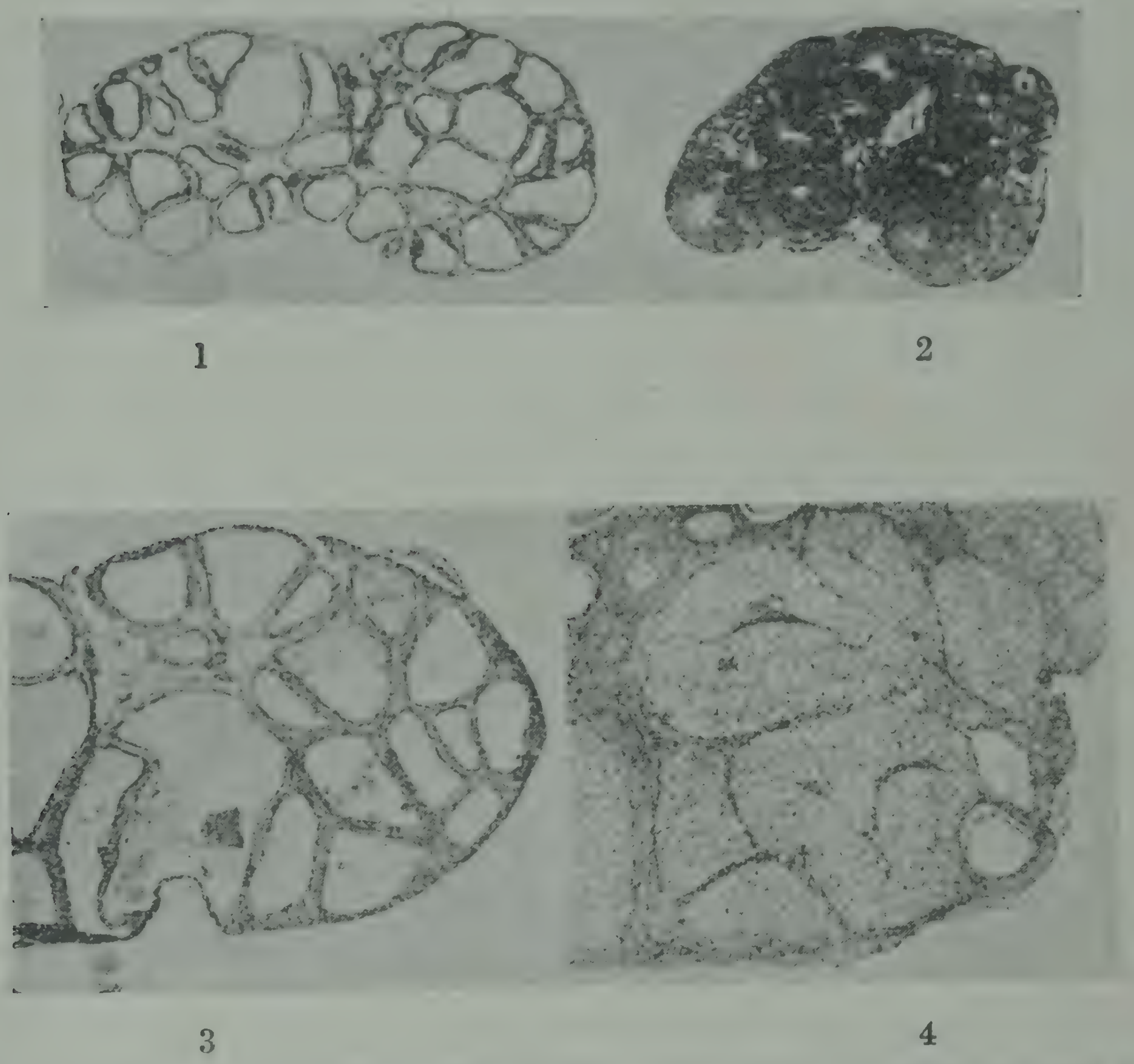


FIG. 666. To illustrate the effect of the follicle-stimulating and luteinising hormones on immature ovaries.

(1) Follicle-stimulating effect on rat, from F.S.H. fraction of 1 g. sheep pituitary. No luteinisation.
(2) Ovary of rat after unfractionated pituitary of sheep, 0·1 g. Full of corpora lutea.
(3) Rabbit, after F.S.H. fraction of 2·5 g. sheep pituitary.
(4) Rabbit, after 2·5 g. whole sheep pituitary.

(After HISAW *et al.*)

The pituitary is capable of stimulation by nerve impulses reaching it from the hypothalamus, *viâ* the stalk.* Fee and Parkes † showed that, in rabbits, removal of the pituitary prevents the normal ovulation following copulation.

---

* BROOKS. *Amer. J. Physiol.*, 1938, **121**, 157.
† FEE and PARKES. *J. Physiol.*, 1929, **67**, 383.

Animals that have an anœstrous season in winter, *e.g.*, ferrets, can be brought into œstrus by prolonged exposure to light. This response to light cannot be brought about in hypophysectomised animals * and is impaired by section of the optic nerves. It is therefore concluded that the awakening of the gonads is brought about by stimulation of the anterior pituitary from the retina. In this reflex, the visual cortex and superior colliculi appear not to be involved, and the path seems to be either to the ventral nucleus of the lateral geniculate body, or to the subthalamus by way of the accessory optic tracts. But how the impulses then reach the anterior lobe is doubtful; it may be *viâ* the hypothalamus,† or it may be humorally conveyed (p. 1094).

There appear to be two "gonadotropic" principles formed in the anterior pituitary (Fig. 666): one, a water-soluble substance, called the *follicle-stimulating hormone* (F.S.H.), stimulating follicular maturation, and one, insoluble in water, called the *interstitial cell-stimulating hormone* (I.C.S.H.), or *luteinising hormone* (L.H.), provoking premature luteinisation ‡; these two hormones are normally excreted in small amounts in the urine. After oöphorectomy, or in states of hypo-functioning ovaries, such as at the menopause, a follicle-stimulating hormone is present in the urine in increased amount. Since in those cases the pituitary body has an increased content of the hormone and also contains numbers of large basophile cells, it is supposed that the F.S.H. is formed in the basophile cells.

Relation of the Corpus Luteum to the cycle. The basal ovarian cycle is modified to a variable extent by the presence of the corpus luteum. This structure has functions specifically connected with pregnancy, apart from which condition its effects are only incidental. These effects are responsible for the *luteal phase* of the œstrous cycle. Thus, effects due to the corpus luteum in the unmated mouse are small, in correspondence with the insignificant development of the corpora formed after ovulation. After mating with sterile (vasectomised) males, the corpora persist for a longer period and attain a greater development; the period of the cycle is then lengthened to about twelve days, during which time the uterus is much enlarged and vascularised, and there is pronounced growth of the mammary glands. This condition is called *pseudo-pregnancy*, and is seen as a normal sequel to the œstrous period in the bitch and other species where the corpora lutea of ovulation persist for a much longer time, irrespective of whether sterile copulation has occurred or not. In the bitch, the uterus is enlarged, with the endometrium thickened and vascularised, and there is enlargement of the mammary glands: as the condition passes off there is breakdown of the endometrium, with hæmorrhage.

In normal pregnancy in all species, where the corpora lutea maintain a still greater development, the œstrous cycle is suspended and all the above phenomena are accentuated. Consideration of this and other facts indicates that the corpus luteum produces a hormone having the following functions:—

(1) Inhibition of œstrus and of ovulation.

(2) Growth and sensitisation of the uterus.

(3) Maintenance of pregnancy by improvement of the nutrition of the uterus.

(4) Enlargement of the mammary glands during pregnancy and pseudo-pregnancy. Also, other special phenomena accompanying pregnancy, *e.g.* the remarkable relaxation of the pelvic ligaments which occurs in the guinea-pig, the mucification of the vaginal mucosa in the rat, modification of the endometrium in the rabbit, etc., etc.

Whether all these effects are to be attributed to the production, in the corpus luteum, of a single hormone, or of more than one, is as yet uncertain.

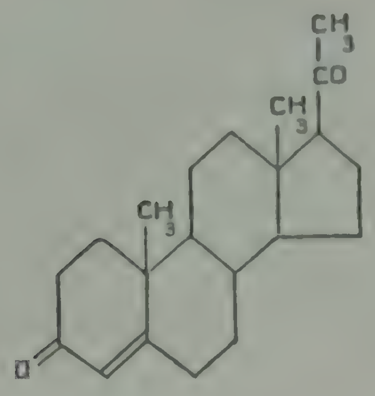
---

* Hill and Parkes. *Proc. R. S.*, 1933, **113B**, 537.

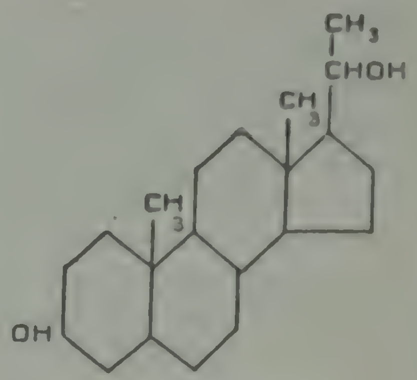
† Le Gros Clark *et al.* *Proc. R. S.*, 1939, **126B**, 449.

‡ Wallen-Laurence. *J. Pharm.*, 1934, **51**, 263; Smith. *J. Amer. Med. Assoc.* 1935, **104**, 553.

One substance has been isolated and identified with certainty.* Its structure is shown below :—



PROGESTERONE



PREGNANDIOL

and it has received the name *progesterone.* If progesterone is administered, it is converted into *pregnandiol*, the glucuronate of which is excreted in the urine.

The inhibition of œstrus and of ovulation which is effected by the agency of the corpus luteum can be shown in a variety of ways. Thus, in the cow the removal of the corpora lutea of ovulation hastens the onset of the next cycle † : conversely, if luteal formation is augmented, as H. M. Evans found to happen after injection of alkaline extracts of anterior pituitary, there is cessation of the

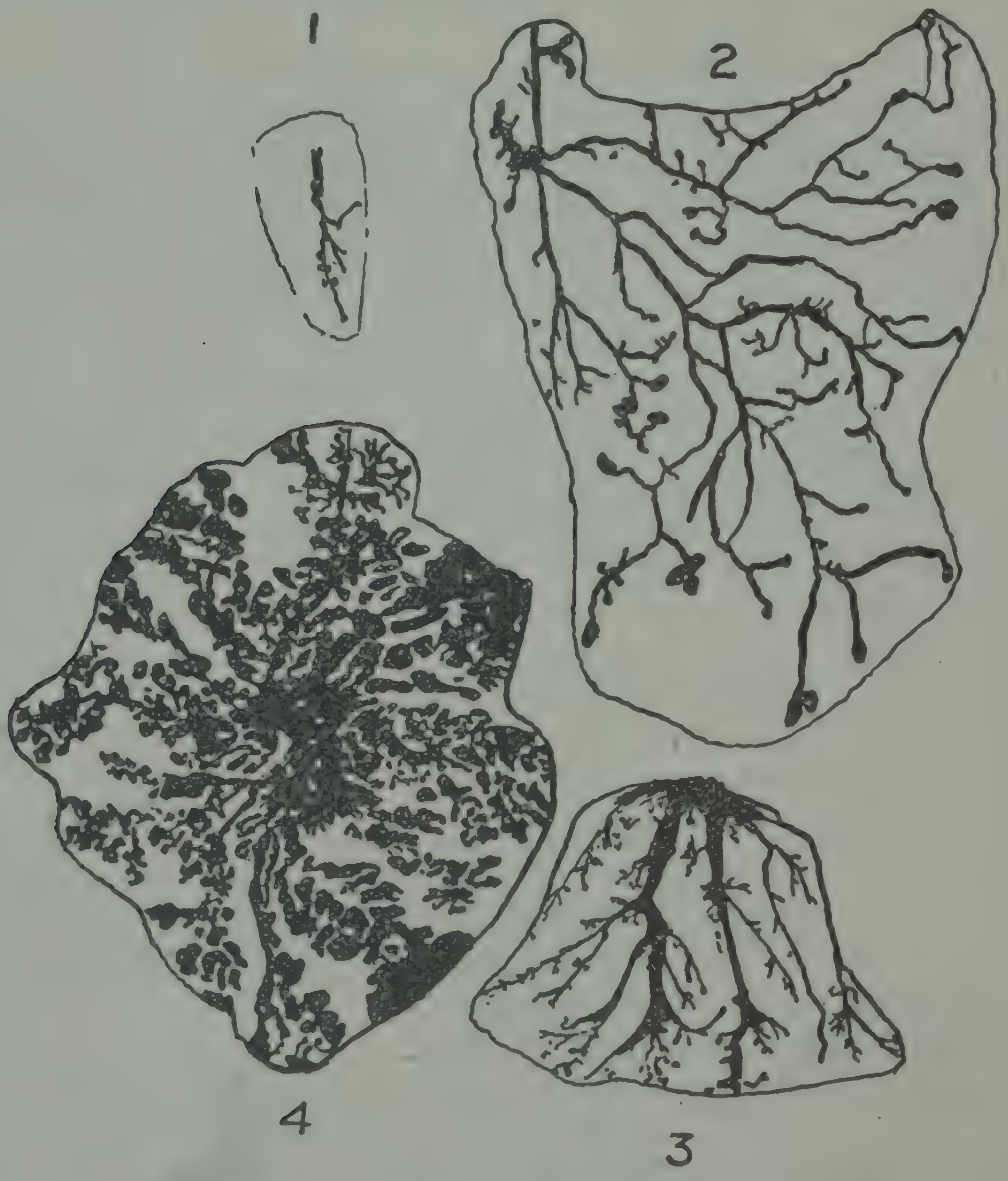


Fig. 667. Development of Mammary Gland by Ovarian Hormones.

(1) Mammary gland of male mouse (× 8).

(2) Mammary gland of male mouse after 4 rat units of œstrogen daily for thirty days. This corresponds to the development in a virgin adult female (× 4½).

(3) Mammary gland of spayed rabbit after twenty days' injections of progesterone, preceded by œstrogen treatment. No corpus luteum type of alveolar growth (× 1).

(4) Similar to (3), but œstrogen and progesterone were given together. Growth and development resembles that at mid-pregnancy. (After Turner, 1932.)

---

* *Annual Rep. Prog. Chem.*, 1934, **31**, 209, 326.

† Hammond. The Physiology of Reproduction in the Cow. 1927. (Cambridge.)

œstrous cycles, together with other changes characteristic of pseudo-pregnancy.

Progesterone appears from the work of Frænkel to be essential for the preparation (sensitisation) of the endometrium for the attachment of the fertilised ovum, at the initiation of pregnancy. It was shown by Loeb that, in the guinea-pig, when corpora lutea were present, mechanical irritation of the uterine mucosa, *e.g.* by a stitch passed through the uterus, led to local accumulation of decidual cells, which represented the attempted formation of a placenta ; but after removal of the corpora lutea, this response vanished. Progesterone facilitates these progestational reactions in rabbits which have been oöphorectomised during œstrus, or which, after oöphorectomy, have received preliminary injections of an œstrogen, but otherwise does not cause them in oöphorectomised rabbits. This means that the uterine response is the result of the joint action of an œstrogen and progesterone.

Other actions of progesterone are an abolition of spontaneous contractions of the uterus, and, when œstrone is also given, a considerable growth of the mammary gland (Fig. 667) with extension of the actual secreting alveoli.

Even after the placenta has been fully formed, however, the corpus luteum appears to be essential for the continuance of pregnancy, and removal of the corpora, or of the entire ovary, is followed in many species by abortion. But if progesterone is administered, oöphorectomy does not terminate pregnancy.

Menstruation. Puberty in the girl is marked by the onset of menstruation. This is a flow of blood and mucus from the uterus, recurring about every 28 ± 3 days and lasting each time from four to eight days ; it is an outcome of a modified œstrous cycle, and occurs in all the Old World primates. In some, including man, the only outward indication is the periodic bleeding ; in others, *e.g.* macaques, baboons, there are areas of skin which become highly coloured during certain phases of the cycle, or at other times if œstrogen is injected.* In normal women, except during pregnancy and lactation, menstrual periods recur at regular intervals until the age of the menopause (45–50) is reached.

Before the first menstrual period, other signs of approaching sexual maturity are usually observed and are attributed to the elaboration, by the ovary, of an internal secretion. These include rapid growth, with changes in the skeleton, leading to the typically feminine pelvis, a development of the genitalia and mammary glands, and the appearance of secondary sex characters, such as the growth of hair on the pubes. At the same time, there is development of the mental characteristics typical of the sex. At the menopause there is some loss of feminine characteristics, and a degree of psychological upset dependent on the suddenness of the change. The amount of blood lost at each menstrual period varies between 10 and 200 grammes. The flow is intermittent, and dependent on uterine contractions ; the blood is not coagulable and contains an anti-coagulant substance. Some toxic substance of unknown nature is said to be present in the sweat during menstruation. During the " period " there are often disturbances of other functions, and a general disinclination for exertion.

CHANGES IN THE UTERUS. In describing the menstrual cycle, it is customary to number the days from the commencement of one flow to the commencement of the next.

During the first (or follicular) half of the cycle, the reproductive system is under the influence of œstrogens formed in the ripening follicles : during the second (or luteal) half under that of both œstrogen and progesterone.

The histological changes† are briefly as follows : at the first day after menstruation the mucosa, or endometrium, is 1 mm. thick and the glands are short simple tubules, between which are fusiform stroma cells. This may be called the *basal zone*. Preparation for menstruation begins at once, and the *phase of regeneration* occupies till the fourth day. It consists in the appearance of a new superficial layer of epithelium, covering up the bare stroma, and about 0·5 mm. thick. Then follows the *phase of proliferation* up to the fifteenth day; in this the superficial layer reaches about 3·5 mm. in thickness, with nuclear division in the gland cells, growth and secretion of the tubules and swelling of stroma cells. The enlargement and secretory activity of gland tubules increases until the twenty-eighth day (*phase of secretion*), when the œdema of the stroma becomes localised to the middle zone of the endometrium, and the capillaries become dilated. From the twentieth to the twenty-eighth day, the progestational changes become more apparent: the superficial zone, with close-packed stroma cells, closely resembles the compact layer of the decidua, while the middle zone, with dilated glands and œdematous stroma, resembles the spongy layer. The total endometrium is now about 6 mm. thick. When the flow begins, the *phase of desquamation* (twenty-eighth to second days), these two outer zones become disorganised, infiltrated with leucocytes and phagocytes, and ultimately are shed, with some bleeding, leaving the basal zone containing the lower ends of the glands, with the stroma bare. After the flow the epithelium from these glands grows over the bare stroma to form a lining (one to four days) and the next cycle has begun (Figs. 668, 669).

---

* PARKES and ZUCKERMAN. *J. Anat.*, 1931, **65,** 272 ; ZUCKERMAN. *B. M. J.*, 1932 (ii.), 1093 ; The Social Life of Monkeys and Apes. 1932. (London : Kegan Paul).

† SHAW. *J. Physiol.*, 1925, **60,** 193 ; BARTELMEZ. *Contrib. to Embryol.*, No. 142, 1933 ; Carn. Inst. Wash. Publ., 435.

Similar histological changes have been seen to occur in small pieces of surviving monkey's endometrium transplated into the anterior chamber of the eye.[*] After preliminary vasoconstriction, subepithelial hæmorrhages occurred and small papillæ formed on the free surface. These ruptured and bled for about half an hour, after which new papillæ formed and ruptured. Desquamation began locally after the first day.

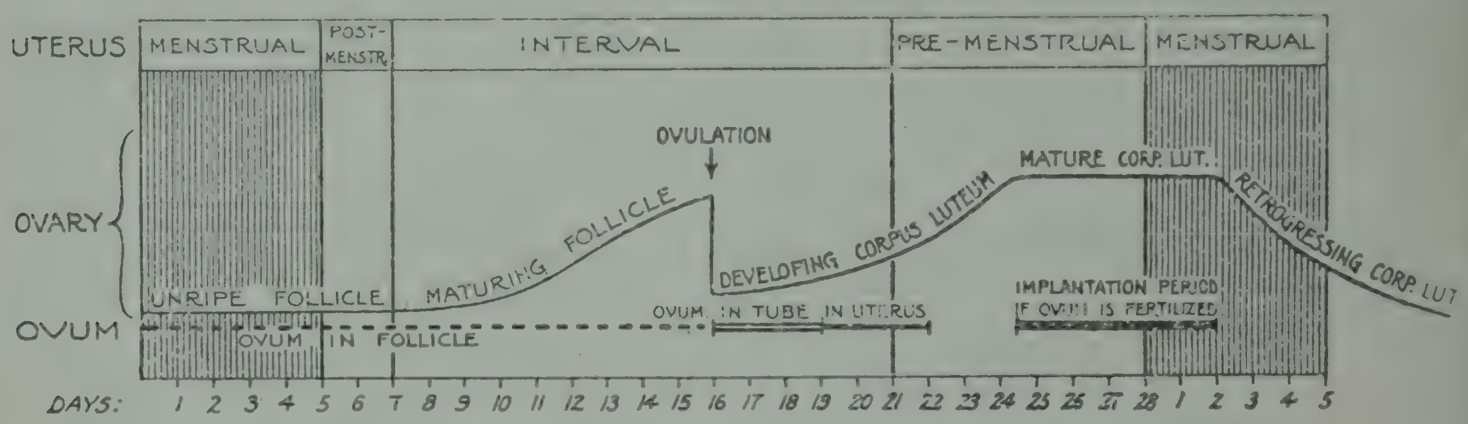


Fig. 668. Course of Events in the Human Menstrual Cycle. (G. W. Corner.)

Rhythmic contractions can be recorded at all stages by suitable small balloons passed into the non-pregnant human uterus.[†] In the first half of the cycle they are small and frequent; in the latter half, stronger but less regular. During menstruation strong waves occur at intervals of two to three minutes (Figs. 670, 671). Vasopressin causes spasmodic contraction of the uterus at all stages of the cycle, but oxytocin has no action at all on the non-pregnant uterus,[†] nor in early pregnancy.[‡]

Changes in the Vagina. The vaginal epithelium and secretions undergo a typical cycle of changes during the menstrual cycle, and these changes can be produced after ovariectomy by administration of œstrogens and progesterone. They are recognised by examination of vaginal smears.[§]

The Causation of Menstruation.[‖] There is no doubt that menstruation normally depends on rhythmical activity of the ovary, in discharging œstrogens and progesterone at appropriate times. This in its turn depends on the output of gonadotropic (follicle-stimulating) hormone from the anterior pituitary. Its onset coincides with the first production of ripe ova in the ovary and it ceases, with the cessation of ovulation, at the menopause. In cases where the ovaries have been removed before puberty, menstruation never occurs. Removal of both ovaries during the reproductive years of life is a serious operation, since it brings about a premature menopause.

Despite its cyclic character the menstrual rhythm cannot be spoken of as "œstrus." The female may experience desire at any period of the cycle, and though coitus may be refused during menstruation itself, or for some time after, this is on æsthetic grounds rather than from absence of desire. Evidence as to the relation of ovulation to menstruation rests chiefly on observations made on the ovaries in abdominal operations (Shaw *) and by Corner, Hartman, and others, on menstruation in monkeys. These tend to show that the whole cycle may be roughly divided into a "follicular" (œstrogen) phase leading up to ovulation, *i.e.* the first fourteen to sixteen days, and a "luteal" (progesterone) phase for the remainder. It seems probable that the ripening and discharge of the ova in the human ovary usually occur about the thirteenth to seventeenth day of the menstrual cycle. At the time of ovulation there are changes in body temperature and, in some women, a characteristic pain ("mittelschmertz") at rupture of the follicle. Live human ova have been found in the Fallopian tubes on the fourteenth day of the cycle,† and the most fertile period for copulation lies in the four or five days before and after mid-cycle. Œstrogen excretion also rises to a peak (600 units daily) at about mid-cycle. In the baboon, the lively colour of the sexual skin, due to processes accompany-

---

[*] Markee. *Anat. Rec.*, 1933, **55**, 66.
[†] Moir. *Edin. Med. J.*, 1934, **41**, 93.
[‡] Robson. *J. Physiol.*, 1933, **79**, 83.
[§] Papanicolaou. *Amer. J. Anat.*, Suppl., 1933, **52**, 519.
[‖] Allen. *J. Amer. Med. Assoc.*, 1935, **104**, 1901.

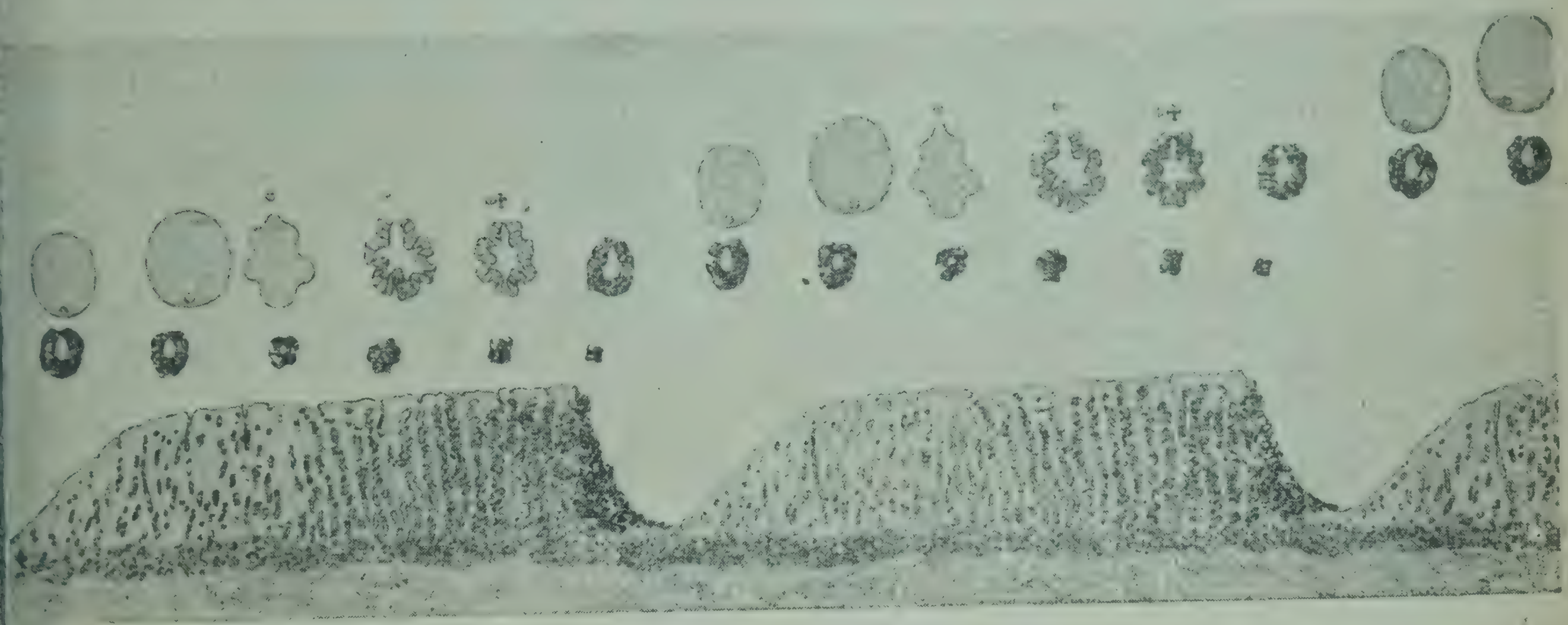


FIG. 669. Diagram showing, from left to right, the Changes in the Human Endometrium (below) and Ovary (top row, this month's follicle or corpora lutea; lower row, last month's) during two complete Menstrual Cycles. (After SCHRÖDER.)

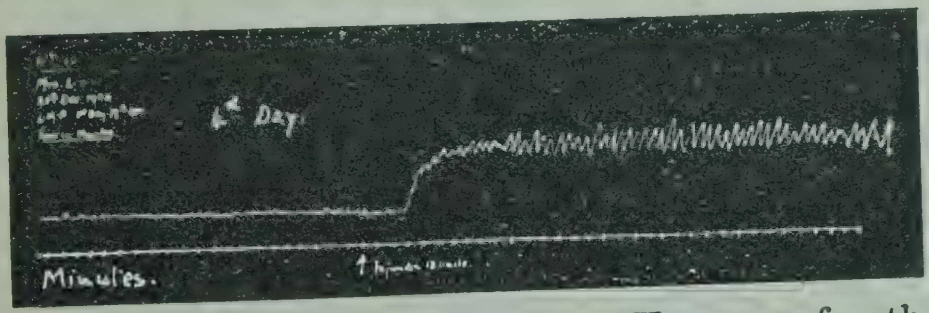


FIG. 670. Contractions of Non-pregnant Human Uterus on fourth day of Cycle, showing small rapid Contractions and Response to 10 Units of Pituitary Extract. (After MOIR.)

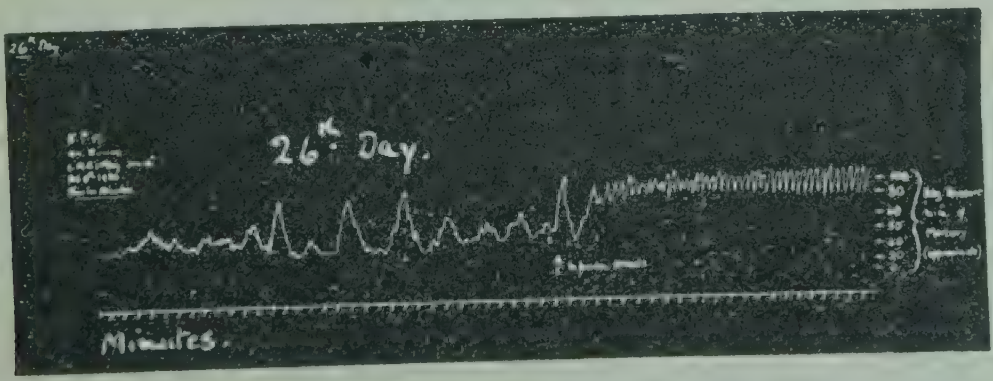


FIG. 671. Larger Contractions of Non-pregnant Human Uterus on twenty-sixth day of Cycle, showing Response to 10 Units of Pituitary Extract. (After MOIR.)

---

* SHAW. *J. Physiol.*, 1925, **60**, 193.

† ALLEN *et al.* *Contrib. to Embryol.*, 1930 (No. 414), **22**, 45.

ing the maturation of the ovum, abruptly subsides at about the mid-point of the cycle and begins to appear again shortly after the onset of menstruation. It can be produced artificially by injection of an œstrogen. During the follicular phase, and as a result of the effect of œstrogen, there is growth of the endometrial glands. According to Frænkel, after ovulation, which normally takes place once in each menstrual cycle, the corpus luteum at once begins to develop and attains maturity in about five days, *i.e.* about the twenty-first day. It persists in this state until the commencement of the menstrual flow, when it usually abruptly degenerates; in any case, it is already undergoing degeneration by the end of the flow. The excretory derivative of progesterone, pregnandiol glucuronidate, is found in the urine from the middle of the cycle (forty-eight hours after ovulation) until just before menstruation begins.* If progesterone is administered to a normal woman she excretes about 45 per cent. of the progesterone in this form. Males also excrete it after dosage with progesterone; the conversion is thus not a specialised metabolic event. In anovulatory cycles no pregnandiol glucuronidate is excreted.

Coincident with the growth of the corpus luteum and the presence of progesterone in the blood, the uterine glands undergo rapid growth and the mucosa undergoes changes preparatory to receiving a fertilised ovum. This, the progestational stage, thus corresponds to the pseudo-pregnancy stage of the lower mammals. If impregnation occurs, the embryo becomes embedded in the thickened membrane, and this continues to grow and to form the placenta, which involves the opening of maternal blood-vessels into large intervillous spaces.

Several theories regarding the significance of menstruation have been advanced. These need not be discussed here, but according to the most popular one it may be regarded as a bid for pregnancy. If this fails to occur, the ovum is lost, the corpus luteum undergoes retrogression, and, possibly in consequence of this, the hypertrophied mucosa breaks down, the *débris* being discharged, together with extravasated blood, as the menstrual fluid. In conformity with this, excision of the corpora lutea at once precipitates a menstrual flow.

It would seem that two, if not three, factors are concerned in the causation of normal menstruational changes, of which the menstrual flow is the final stage. For the proper growth changes in the endometrium, œstrogen and progesterone appear to be essential; for the causation of the menstrual flow, withdrawal or reduction in both hormones is probably necessary.† First, there is the follicular phase in which the uterus and the endometrial glands enlarge and become highly vascular. This, the proliferative phase, can be imitated in oöphorectomised monkeys by the injection of large doses of œstrogen. Stoppage or great reduction of the dose of œstrogen leads to a menstrual flow in about a week. Secondly, there is the luteal phase, in which the endometrium undergoes the characteristic changes. This phase can be imitated in spayed monkeys by following up the œstrogen injections by injections of progesterone, beginning at the fourteenth day, when the œstrogen dose is reduced. Lastly, there appears to be a phase which is responsible for the actual necrosis of the prepared endometrium; this may be imitated, in the experiment quoted above, by stopping the progesterone after ten days, *i.e.* at the twenty-fourth day, as a result of which bleeding begins about the twenty-eighth day.

A difficulty which cannot be ignored in all theories of menstruation is the unquestionable fact that normal menstruation without ovulation or corpus luteum formation often occurs in apes, and sometimes in women. In this case we might assume that there have been no progestational changes in the endometrium, and that menstruation is due merely to cessation of œstrogen production.

---

* Venning and Browne. *Endocrinol.*, 1937, **21**, 711.

† Parkes. *B.M.J.*, 1931 (ii.), 559.

The probable course of events in the human menstrual cycle is summarised in the diagram on p. 1116 (Fig. 668). The changes in the endometrium and in the ovary are shown diagrammatically in Fig. 669.

The Menopause. At the menopause, ovarian function ceases. The condition normally appears more or less gradually, and the symptoms vary in severity, but it may be artificially induced by surgical removal of both ovaries, or by destruction of the ovaries by deep X-ray or radium treatment, and then the symptoms are more severe.

Cessation of menstruation may be sudden, or may be preceded by irregularity. The breasts, uterus and external genitalia atrophy and sexual desire ultimately wanes. These phenomena are to be attributed to the reduction or loss of ovarian hormones. Aspects of virilism may become apparent, such as growth of hair on the face, and an alteration of the features, and these have been said to be due to overaction of the suprarenal cortex. The anterior pituitary may become overactive too, as shown by an increase in follicle-stimulating hormone in the urine, or even by signs of incipient acromegaly. The thyroid activity is lessened, and definite hypothyroid symptoms are often encountered.

Changes in the circulation, such as hot flushes or palpitation are common. Perhaps most striking, though most often concealed, are mental and emotional instability, which may be long-lasting, or even go on to clear-cut insanity.

For Reference

Allen *et al.* (1939). Sex and Internal Secretions. (London : Baillière, Tindall & Cox.)

Bartelmez (1937). Menstruation. *Physiol. Rev.*, **17**, 28.

Corner (1938). The Sites of Formation of Estrogenic Substances in the Animal Body. *Physiol. Rev.*, **18**, 154.

Gardner and Pfeiffer (1943). Influence of Estrogens and Androgens on the Skeletal System. *Physiol. Rev.*, **23**, 139.

Hartman (1936). The Time of Ovulation in Women. (Baltimore, Williams & Wilkins.)

Hisaw (1947). Development of the Graafian follicle and Ovulation. *Physiol. Rev.*, **27**, 95.

Laqueur (1945). Inter-relationship between Gonadotrophic and Sex Hormones. *Harvey Sects.*, **41**, 216.

Marrian (1933). Recent Advances in the Chemistry and Biological Assay of Œstrin, *Physiol. Rev.*, **13**, 185.

Marshall (1942). Exteroceptive Factors in Sexual Periodicity. *Biol. Rev.*, **17**, 68.

Parkes (1945). The Adrenal-Gonad Relationship. *Physiol. Rev.*, **25**, 203.

Reynolds (1939). Physiology of the Uterus. (London : Hamilton.)

Speert (1948). Local Action of Sex Hormones. *Physiol. Rev.*, **28**, 23.

THE MALE REPRODUCTIVE ORGANS

In all the higher animals, we may divide the male reproductive organs into the essential organs, which form the germ cells, the spermatozoa, and the accessory organs, which have as their office the facilitation of the access of the spermatozoa to the ova (fertilisation). We also have to consider male secondary sexual characters, *e.g.*, the comb of the fowl, the horns of the stag, and, in man, the body build and stance, the deep voice, and the specialised distribution of hair on face, chest and pubic region.

The essential sexual organ of the male is the testis. This is made up of a collection of convoluted tubules, the seminal tubules, which are contained in a number of compartments separated by fibrous septa. The tubules present few or no branches, each one being about 500 mm. long. Several tubules unite to form a straight tubule, which leads by a series of communicating spaces, the rete testis, into the vasa efferentia (Fig. 672). These join to form the duct of the epididymis, coiled into a mass lying at the back of the testis. The epididymis is composed of the convolutions of this single duct, which is about 20 feet long. From the lower end of the epididymis, the vas deferens, a tube with thick muscular walls, leads by the abdominal ring to the base of the bladder, where it opens into the beginning of the urethra in its prostatic part. Just before it joins the urethra as the ejaculatory duct, each vas deferens is joined by a diverticulum, the seminal vesicle, which lies along, and is attached to, the base of the bladder. The seminal vesicles are lined by columnar epithelium, and have plain muscle in their walls. They contain a glairy liquid which may, but which usually does not, contain some spermatozoa. The prostate, which surrounds the first part of the male urethra, is divided into lobules by septa of unstriated muscle-fibres, enclosing numerous wide branched racemo-

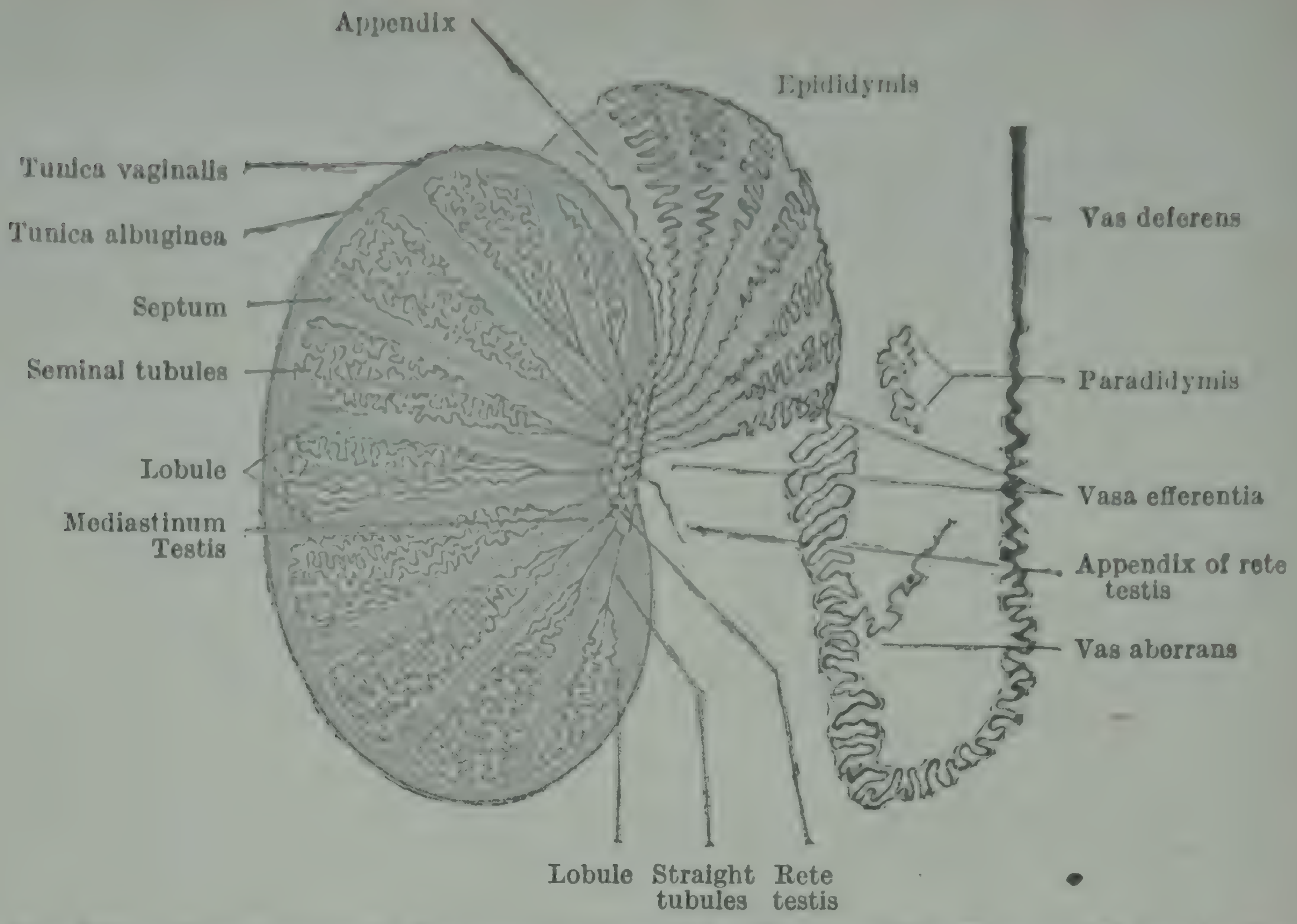


FIG. 672. Diagrammatic Representation of the Course of the Seminal Tubes in the Testis and Epididymis. (After NAGEL).

tubular glands. These open by twenty to thirty ducts into the floor of the prostatic urethra, near the entry of the ejaculatory ducts. A blind recess from the prostatic urethra, called the utriculus, is a vestige of the Müllerian duct, which, in the female, forms the Fallopian tubes, uterus and vagina. In some females, para-urethral glands homologous with the prostate are found. From the point of entry of the vasa deferentia to its orifice, the urethra represents a common passage for the urine and semen. It passes, therefore, through tissues forming the penis, which is especially adapted for the introduction of the semen into the female. In the urethra we distinguish the prostatic, the membranous and the penile portions. Into the beginning of the penile portion, the bulb of the urethra, open the ducts of the two small racemose glands of Cowper. In the penis itself, the urethra is surrounded with erectile tissue, forming the corpus spongiosum, and lies between the two corpora cavernosa, which consist of the same kind of tissue. The erectile tissue is a spongy meshwork of elastic and unstriated muscle fibres, enclosing spaces in free communication with the efferent veins of the organ. The arterioles also open into these spaces, but under normal circumstances both the arterioles and the muscle tissues of the framework are contracted, so that the blood trickles very slowly from the arterioles into the spaces, whence it escapes readily by means of the veins. If these muscle fibres be relaxed, so that the blood can pass rapidly into, and distend, the spaces, the tissue swells and becomes harder, causing " erection " of the organ.

Hyaluronidase. Extracts of testis contain an enzyme, which has the properties of a " *diffusion factor.*" * Mixed with any coloured substance, the extract greatly augments the speed and extent of diffusion of the dye, when injected intracutaneously or subcutaneously. The property is due to the presence of an enzyme, hyaluronidase, which hydrolyses the mucin-like interfibrillar substance in the dermis.† The enzyme is present in spermatozoa.

Puberty and Spermatogenesis. In the immature testis, *i.e.* from birth up to puberty, the seminal tubules are lined with cells having large nuclei. Some of these are the spermatogonia, the mother cells of the future spermatozoa, while the others form the cells of Sertoli, whose function it is to act as nurse cells to the developing spermatozoa. In hypophysectomised animals the testes remain in, or revert to, an immature condition. The actual formation of spermatozoa begins at puberty, and can be produced precociously (or in a hypophysectomised animal) by the " follicle-stimulating " hormone of the anterior pituitary (Fig. 673) which, as we have seen, is also responsible for the maturation of ova in the female. Since this hormone has a stimulating action on the production of gametes of both sexes, it is therefore often known, perhaps more appropriately, as the *gametokinetic hormone*; other pituitary hormones cause hypertrophy of the interstitial cells of the testis, and this tissue then produces male sex hormones which cause growth and development of the various accessory organs of reproduction.

---

* McCLEAN. *J. Path. Bact.*, 1930, **33**, 1045.
† CHAIN and DUTHIE. *Nature*, 1939, **144**, 977.

In the formation of spermatozoa, the spermatogonia divide many times to form the spermatocytes, which in their turn undergo reduction division to form the spermatids, as already described. By a further division and modification of the latter, the fully developed spermatozoa are formed. These, when mature, pass by the tubules of the testis into the epididymis, where they are stored until an ejaculation takes place, or failing that, die and are reabsorbed. Their movement is probably facilitated by the ciliated cells lining the tubule of the epididymis, as well as by the secretion of the lining membrane of the seminal vesicles. The spermatozoa are motionless while in the seminiferous tubules of the testis, but become actively motile in the vas deferens, or when mixed with prostatic secretion. They can then travel, head foremost, at

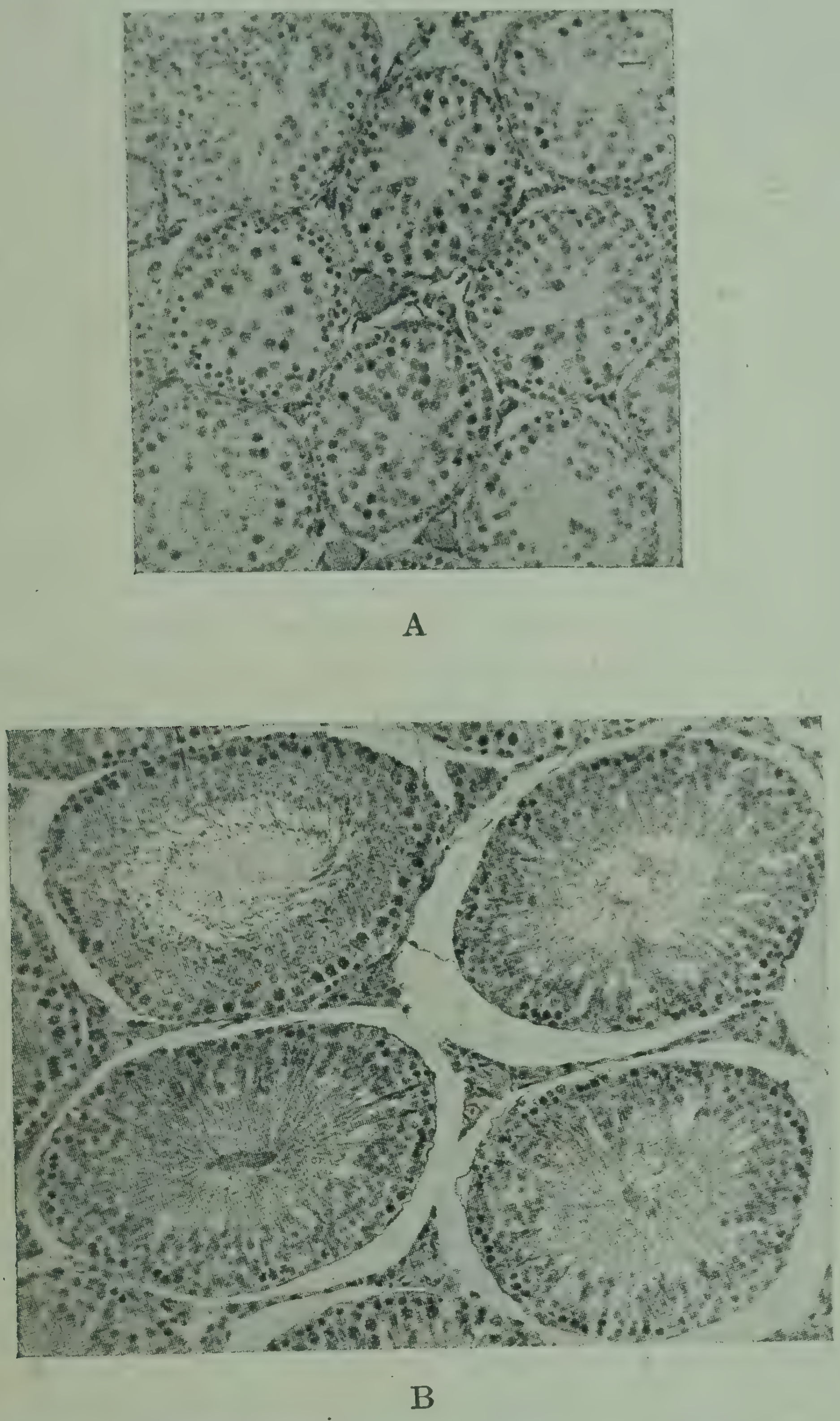


FIG. 673. Testis of Hypophysectomised Rats. A. After twenty days. No treatment. Tubules atropic. B. Testis of litter mate of above animal, treated for twenty days by injection of pregnancy urine gonadotrophin. The animal was mated, and fertile, on the tenth and twentieth days. (After SMITH and LEONARD.)

about 20μ per second (70 mm. per hour). Spermatozoa recovered from the female tract soon after their introduction are always in active motion. This is soon lost, however, in most species, including man (probably in thirty to forty hours), and fertilising power is lost sooner still. Spermatozoa survive longest at *p*H.7·0 and low temperatures ; on being warmed to body temperature the activity is increased, but the length of survival shortened. The energy is derived mainly from glycolysis, oxygen usage being slight. The formation of the spermatozoa is continuous, though the rate at which this occurs is variable and is regulated by the sexual activity of the individual. In many mammals, including man, the optimal temperature for spermatogenesis is below the normal body temperature ; hence if the testis fails to descend into the scrotum, where a lower temperature is maintained, it becomes non-functional.

The Prostate is an important accessory organ of reproduction, inasmuch as its secretion acts as a vehicle of proper bulk and composition for the suspension of the spermatozoa. Theoretically, it is not indispensable for fertilisation, which, in some species, has been effected by artificial insemination with the contents of the epididymis. But without it the volume of the sperm would be so greatly reduced, and its viscosity so much increased, that fertility, as commonly understood, would probably be abolished. Though other claims have been made, there is no convincing evidence that it has other functions, such as that of furnishing an internal secretion.

The size and activity of the prostate is much influenced by the presence of the sex hormones. Normally, the active secreting cells of the acini are tall—about 25μ. After castration, the prostate undergoes involution, the cells lose in height, and their secretory activity declines ; but if the castrated animal is then given daily injections of an androgen, these effects are reversed, and even hypertrophy ensues. An œstrogen exerts the opposite effect, and in suitable doses can annul the effects of an androgen. Stilbœstrol is given for this reason in cancer of the prostate.

In the rat, but not in man, the secretion of the anterior lobe of the prostate causes coagulation of the secretion of the seminal vesicles.

In man, as age advances, the peri-urethral portion of the prostate hypertrophies and develops characteristic nodules ; if this growth is excessive it may obstruct the urethra and necessitate surgical intervention. The cause of this enlargement is not known ; the cells are of the tall type, suggesting active androgenic effects, to which perhaps the senile gland is hyper-sensitive.

The prostate forms a small amount of " resting " secretion, about 1 c.c. per day in man, and this can be collected from the meatus during massage of the prostate *per rectum ;* from dogs, after an operation by which the urine flow is diverted, much larger amounts can be collected,* up to four times the weight of the gland per hour.

The rate of secretion is increased in sexual excitement, by the stimulation of the nervi erigentes or hypogastrics, and under the influence of androgens, or of pilocarpine ; it is diminished or abolished by œstrogens.

The prostatic secretion has an acid *p*H, and is characterised by the presence of large amounts of an enzyme, acid phosphatase, with an optimum *p*H of 5–6. This enzyme is also present in abundance in the gland itself, and gets into the urine with its secretion. Other notable chemical properties are the presence of proteoses, or other proteins of low molecular weight, of very little glucose, of a high content of citrate and of spermine, and of lipides consisting of cephalin and cholesterol, often in aggregates.

The secretion arouses the spermatozoa to active movement, but the reason for this is not known.

The Semen. In the fully formed semen, the spermatozoa are mixed with the fluids secreted by the lining membrane of the epididymis, the semina vesicle, the prostate and of Cowper's glands. Nevertheless, it contains spermatozoa in enormous numbers, the semen emitted at a single act of coitus (2 to 4 c.c.) usually containing 200 to 300 million spermatozoa.* Though the majority of these are possibly capable of fertilising an ovum, this act is carried out by only one. There is no relation between numbers of spermatozoa and volume of ejaculate, nor between spermatozoa motility and age of the man. Frequent ejaculations may considerably reduce the numbers of spermatozoa in the semen.

---

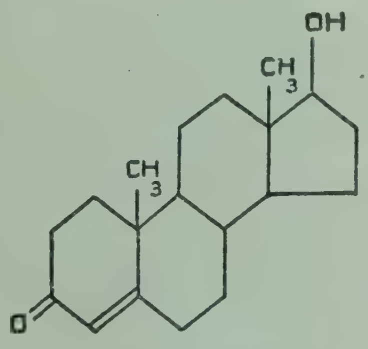
* Huggins *et al.* *J. Exp. Med.*, 1939, **70**, 543.

It is probable that in ejaculation the secretion of the Cowper's glands, with which the urethra is already filled, is expelled first; then follows the glairy prostatic secretion mixed with increasing numbers of spermatozoa and finally the viscous secretion of the seminal vesicles. In the mouse and some other animals this last secretion is coagulated by a vaginal enzyme to form a solid vaginal plug which prevents loss of semen. Human semen also clots,† but is afterwards liquefied again, by the action of a proteolytic enzyme.‡

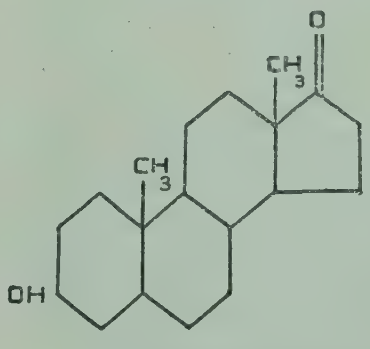
For anaerobic survival a glycolysable carbohydrate, *e.g.* glucose, fructose or mannose is essential. Semen actually, in some species, contains fructose in concentrations up to 1 per cent.§ It is formed in the seminal vesicles, but not after castration.|| Citric acid is another constituent of semen in some species,¶ and is formed in the accessory glands, chiefly seminal vesicles.

THE MALE HORMONES. Between the seminiferous tubules are groups of so-called *interstitial cells,* which appear to have an important relation to the sexual life of the individual. Removal of the testes before puberty prevents the development of the accessory organs, prostate, seminal vesicles, Cowper's glands and penis, and of the secondary sex characters; ligature of the vasa deferentia, which is followed by complete degeneration of the seminiferous tubules, but not of the interstitial cells, does not have these consequences. Hence it is assumed that the development of these characters is to be ascribed to the formation, by the interstitial cells, of an internal secretion having that influence on the individual. This is supported by the observation that when testis grafts are made into castrated animals, the male secondary sex characters develop in the usual way. Masculinisation of ovariectomised females has also been claimed to occur when testis substance is grafted.

Androgens. An androgen is a substance which causes the development of male organs and characters. It is a parallel in the male of the œstrogen in the female. A hormone having these effects has been extracted from testes; it is called *Testosterone,* and a similar substance, called *Androsterone,* is excreted in small amounts in the urine, of both sexes. In male urine, from which androsterone was isolated by Butenandt,** Schoeller and others, the total daily output of all 17-ketosteroids together is about 9 mg.: in female urine it is about 6 to 7 mg. In adrenal tumours the amount is much increased, *e.g.* it may be over 200 mg. daily. The crystalline materials are hydroxy-ketones, insoluble in water, which have been made synthetically and are closely related



TESTOSTERONE



ANDROSTERONE

---

* Hotchkiss and Brunner. *Amer. J. Med. Sci.*, 1938, **196,** 362.

† Goldblatt. *Biochem. J.*, 1935, **29,** 1346.

‡ Huggins and Neal. *J. Exp. Med.*, 1942, **76,** 527.

§ Mann. *Nature,* 1946, **157,** 79.

|| Mann and Parsons. *Nature,* 1947, **160,** 294.

¶ Humphrey and Mann. *Biochem. J.*, 1949, **44,** 97.

** Butenandt. *Z. f. angew. Chem.*, 1931, **44,** 905.

to the female œstrogens.* Testosterone is usually administered as the propionate, subcutaneously injected, in sesame oil. Many other substances also have androgenic activity. It should be noted that there is no sharp line of demarcation between androgens and œstrogens, each being often capable of producing the effects of the other if given in large doses. Yet the effects of an androgen are in many ways antagonistic to those of an œstrogen (Figs. 674, 675). Among the physiological effects produced is a growth of the comb in

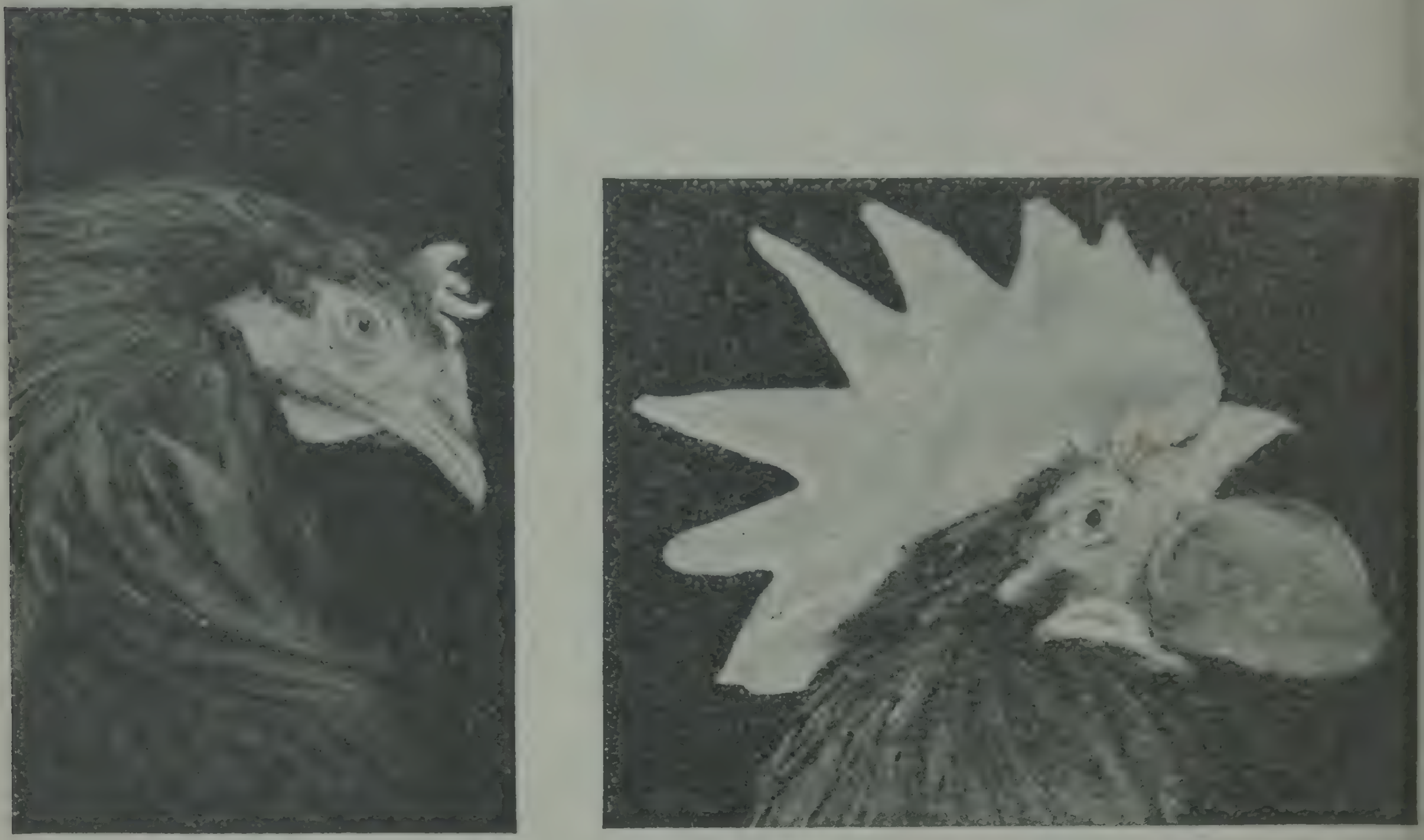


FIG. 674. Effect of an Androgen on a Capon's Comb. On left, the capon. On right, the same after twenty-two daily injections of androsterone. (After PARKES, *Bioch. J.*, 1935, **29**, 1422.)

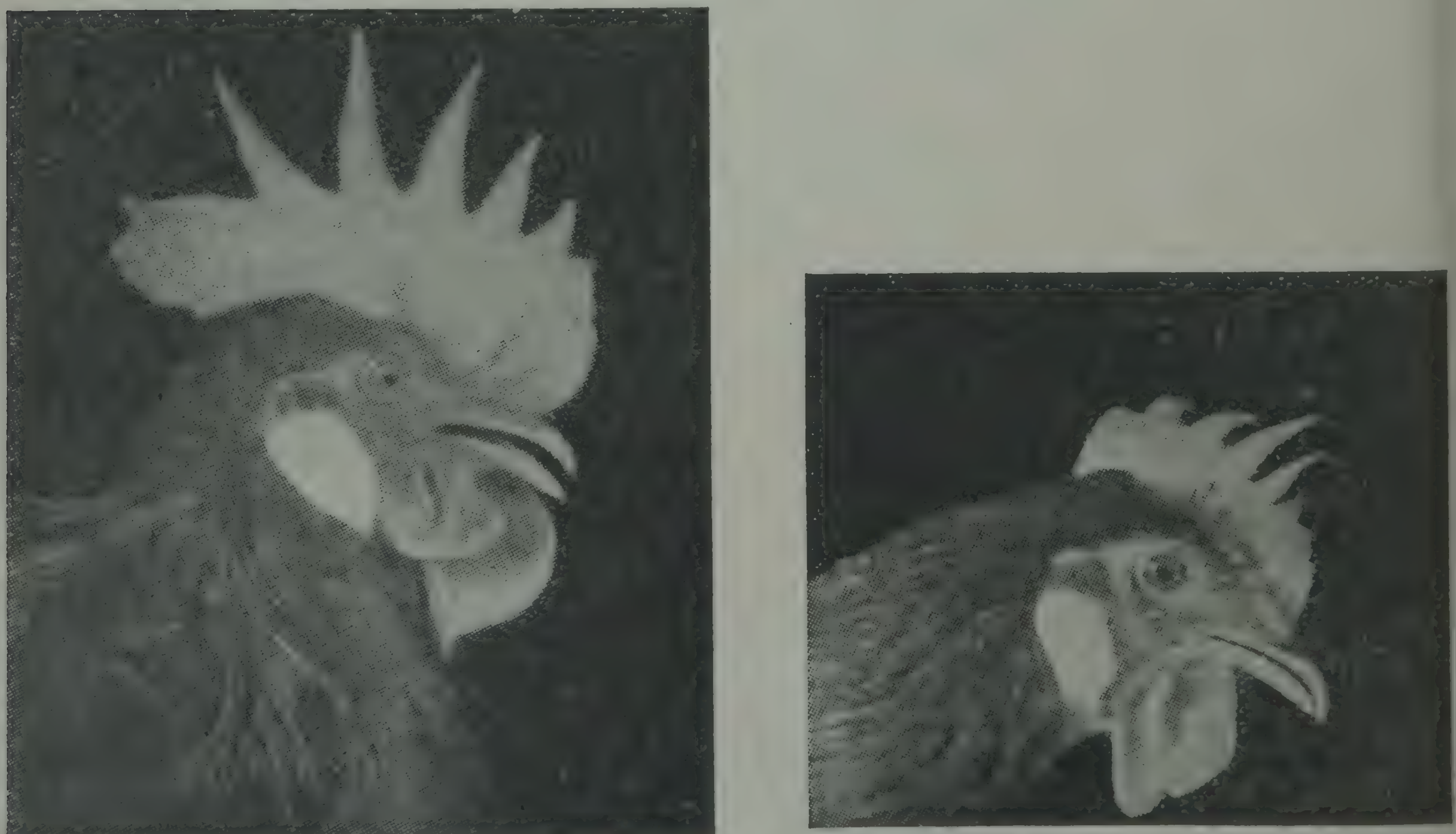


FIG. 675. Effect of an Œstrogen on a Cock. *Left*, before, and *right*, after implantation of tablet of 65 mg. of œstrone. (After EMMENS, *J. Physiol.*, 1939, **95**, 386.)

---

* *Ann. Rep. Progress of Chem.*, 1934, **31**, 323.

capons * and it also prevents the atrophy of the prostate, seminal vesicles (Figs. 676, 677), Cowper's glands, vas deferens, etc., of castrated rats.† Assay of androgens can be made by measurement of the growth caused in the comb of the capon (castrated cock) (Fig. 674), or by the increase in weight of the prostate and seminal vesicles of immature rats (Fig. 676). The international unit is 0·1 mg. of androsterone.

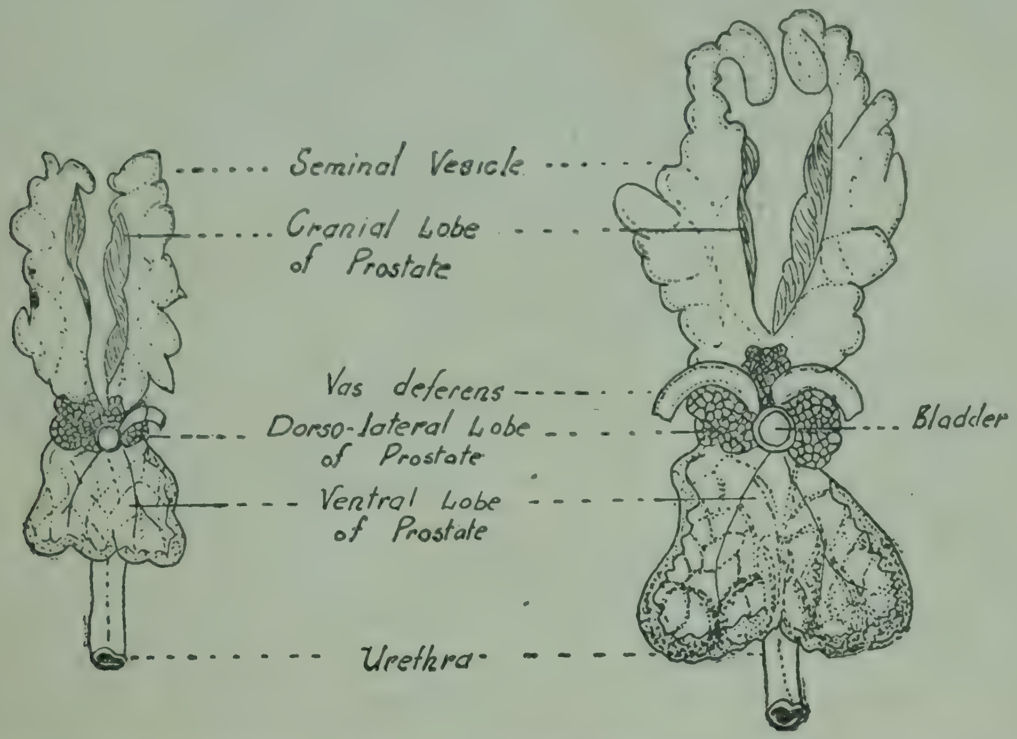


Fig. 676. Internal Genitals of two young Rats of same Litter, showing effect of Androgens. (Drawn by G. Brownlow.) (Gaddum. "Pharmacology", Oxford Press, 1940.)

In man and various other "constant breeders", *e.g.* rats and guinea-pigs, the hormone is constantly produced from puberty onwards throughout sexual life ; it is not stored, but regularly escapes in the urine. Hence atrophy of the prostate, etc., commences in a few days if castration is performed. The rate at which the hormone is produced depends on the state of nutrition and general health, but it ceases at once if the anterior lobe of the pituitary body be removed. This indicates that the anterior pituitary provides a hormone which provokes

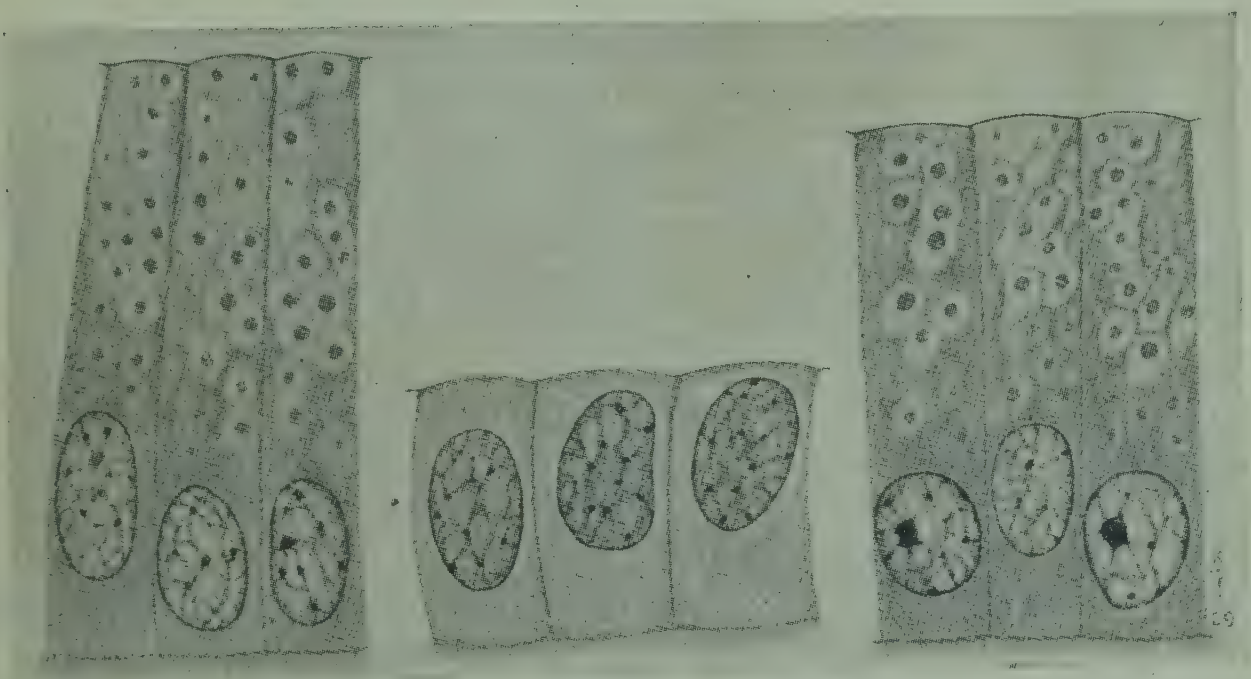


Fig. 677. Effect of Castration and of Androgen on the Epithelium of the Seminal Vesicles. (1) Cells from normal animal, showing secretion granules ; (2) cells from twenty-day castrate ; (3) cells of twenty-day castrate treated with an androgen. (After Moore, Hughes and Gallagher.)

---

* McGee *et al.* *Amer. J. Physiol.*, 1928, **87**, 406 ; Blyth, Dodds and Gallimore. *J. Physiol.*, 1931, **73**, 136.

† Moore. *J. Amer. Med. Assoc.*, 1935, **104**, 1405.

the production of androgens, and further proof of this is shown by the fact that the testicular function may be restored to hypophysectomised males by injection of fresh anterior pituitary or of the F.S.H. (gametokinetic) fraction. Conversely, androgens, and perhaps other substances of testicular origin, can inhibit the production of the gametokinetic hormone by the pituitary.

In those animals such as squirrels, which show seasonal sexual activity, the anterior pituitary seems to be non-functional as regards its gonadotropic functions during the resting phases, and this is probably the case also in all males prior to puberty.

A female hormone, or *œstrogen*, is also present in male urine. This hormone, when injected into males, leads, in some species, to epithelial growth and even cornification in the utriculus, together with diminution of prostatic gland substance and increase in its plain muscle stromata.* It should be noted that the androgens and œstrogens found in the urine of both sexes do not entirely disappear after gonadectomy: they are therefore believed to be, in part, produced in the adrenals.

For Reference

Huggins (1945). The Physiology of the Prostate Gland. *Physiol. Rev.*, **25,** 281.
Huggins (1946). The Prostatic Secretion. *Harvey Lects.*, 1946–47, 148.
Koch (1938). The Chemistry and Biology of Male Sex Hormones. *Harvey Lects.*, 1937–38, p. 205.
Koch (1937). The Male Sex Hormones. *Physiol. Rev.*, **17,** 153.
MacCallum (1937). Pathological Physiology of the Prostate. *Physiol. Rev.*, **17,** 73.
Meyer (1947). The Biological Significance of Hyaluronic Acid and Hyaluronidase. *Physiol. Rev.*, **27,** 335.

DEGREE OF SEX-SPECIFICITY OF HORMONES

The sex hormones are not completely sex-specific, *i.e.* the hormones which characterise one sex are not absent from the other, nor without action on it. In the case of the pituitary hormone, we have seen that the substance which stimulates the growth and development of the gonads is probably the same in both sexes, and for that reason is sometimes called the gametokinetic hormone. The gonadotropic hormones of pregnancy urine (P.U.), of pregnant mare's serum (P.M.S.) and of castrate urine (C.U.), not only cause follicle stimulation or luteinisation, but also have effects on the testis, causing it to increase in weight, by growth of the interstitial cells (by P.U. and C.U.) or of the seminiferous tubules (by P.M.S.).

As regards the œstrogens and androgens, substances of both types are present in the urines of both sexes, though usually the androgen predominates in male urine and the œstrogen in female urine. Moreover, androgens when injected into females may produce two kinds of effect. One of these is a tendency to masculinisation, *e.g.* there is a growth of the clitoris, a masculine type of growth, and some of the male secondary characters may appear, *e.g.* growth of hair on the face in women. The female cycle may also be suppressed, perhaps as a result of an inhibitory effect on the pituitary gonadotropic hormone. The other effect may be œstrogenic in nature, *e.g.* growth of the uterus, or duct development of the mammary gland.

The administration of œstrogens to males, especially to castrates, may similarly produce feminisation, as seen in growth of the mammary glands and nipples, and growth of the prostate, utriculus masculinus, and seminal vesicles, accompanied by cessation of the secretory activity of these organs. But androgenic effects can also be obtained ; these are best seen in the reinforcement of the effects of testosterone when given together with an œstrogen. Progesterone also has effects on the development of the male mammary glands, which resemble those on the female.

The explanation of these common properties is no doubt to be sought in the chemical relationship between these hormones, and is confirmed by the fact that suprarenal cortical hormones, which are also nearly related chemically, can in large doses produce both androgenic and œstrogenic effects.

---

* Parkes and Zuckerman. *Lancet*, 1935, **228,** 924.

VITAMINS IN RELATION TO REPRODUCTION

Animals in a poor state of nutrition do not reproduce readily, so that in conditions of vitamin lack some faults of reproduction are not unusual. Two of the vitamins, however, seem to be especially concerned, viz. vitamin E and vitamin A.

Vitamin E (*v.* p. 824) is essential in the diet of some of the lower animals, *e.g.* rats, if reproduction is to occur. Lack of it, although not apparently interfering with the normal reproductive processes of the female up to the eighth day of gestation, ultimately leads to retarded growth of the fœtus, failure of proper development of the vascular system and allantois, and ultimate death of the fœtus. The maternal placenta develops normally. The vitamin can apparently be stored in the body for several weeks, *i.e.* actual administration of it during pregnancy is not essential, provided that plenty has been given beforehand.

In the male rat, lack of vitamin E leads to atrophy of the seminiferous tubules.

Vitamin A also plays some part in the proper maturation of the gonads, since lack of it leads to abnormalities in the œstrous cycle, and delay in reaching maturity.

COPULATION AND IMPREGNATION

The spermatozoa, stored in the epididymis and vas, are introduced into the female genital tract by the act of copulation, or coitus. After their entry they may ultimately come in contact with, and fertilise, the ovum, discharged from the ovary by bursting of a Graafian follicle. Before coitus can occur erection of the penis must take place. The mechanism of erection is twofold. The most important factor is an active dilatation of the vessels of the penis, especially of the smaller arteries. If the penis be cut across while in the flaccid condition, venous blood merely trickles away from the cut surface, whereas, if erection be excited, the flow of blood from the cut surface is increased eight to ten times, and the blood becomes bright scarlet in colour. A second factor is the contraction of the *ischio-cavernosus* or *erector penis* muscles, certain fibres of which pass over the dorsal veins of the penis and compress these vessels when they contract. Since ligature of the veins coming from the penis does not produce erection, the contraction of this muscle must be regarded as simply aiding the effects of the arterial dilatation. Hence it is possible to excite erection after the second factor has been abolished by paralysing the muscles by curare.

Before, or at the beginning of, coitus analogous changes occur in the female organs, leading to erection of the clitoris and of the erectile structures of the vulva. The glands of the vulva, especially the glands of Bartholin, secrete a mucous fluid, thus lubricating the passage into the vagina. The friction between the glans penis and the wall of the vagina ultimately causes a reflex discharge of motor impulses in both male and female (the "orgasm"). In the former, the muscular walls of the vasa deferentia and seminal vesicles enter into rhythmic contractions, thus ejaculating the spermatozoa and secretions they contain into the urethra. The spermatozoa, mixed with the secretions of the epididymis, the seminal vesicles, the prostatic glands, and the glands of Cowper, form the semen, which, at the orgasm, is ejaculated from the urethra by rhythmical contractions, from behind forwards, of the bulbo- and ischio-cavernosi muscles.

The exact physical events of the female orgasm are uncertain. It has been stated that movements take place coincidently in the uterus, so that its axis more nearly corresponds to that of the vagina, and that the os uteri dilates. After coitus the uterine contractions are diminished for some hours.* The female orgasm, however, does not play an essential part in impregnation, which, as shown by Spallanzani in the 18th century, can be effected by artificial insemination, *i.e.* injection of semen into the vagina by a syringe. It is improbable that any material is aspirated into the uterus during coitus, as was at one time supposed. A more important factor is probably the movement of the spermatozoa themselves.† As we have already seen, these are introduced into the female passage in countless numbers. They will be chemiotactically attracted by the alkaline mucus, secreted by, and filling, the cervix of the uterus. Spermatozoa move at the rate of 1 to 3 mm. per minute. Thus, on a straight course, they might traverse the distance of 16 to 20 cm. between the os uteri and the trumpet-shaped orifice of the Fallopian tubes in three-quarters of an hour. But direct injection into the uterus would shorten this time greatly, and in animals spermatozoa have been found at the peritoneal end of the Fallopian tubes within fifteen minutes after coitus. In women, spermatozoa can be found in the uterus in three minutes. Although ripe sperms remain viable for weeks within the epididymis, their life is brief once they have been introduced into the female tract, where in man few survive more than a day or two. Fertilisation of the ovum normally occurs in the Fallopian tube, and the fertilised ovum is then carried slowly down the tube into the uterus, partly by peristaltic contractions, partly by ciliary action of the tubal mucosa. If not fertilised within a few hours of ovulation, the ovum dies.

Although fertilisation is effected by a single spermatozoon, very many more must be present before it is likely to occur. The minimum number required varies according to their vigour, to species, etc., but in the rabbit, for instance, fertilisation fails if the number is less than about 10,000.‡ In man, sterility is usual if the spermatozoa are fewer than 20,000,000 per c.c.,§ *i.e.* about 60,000,000 in an ejaculate, which would normally contain about four times that number. The explanation of this strange state of affairs is that the ovum, as discharged, is imbedded in the *discus proligerus*, and so long as this remains, the spermatozoa cannot penetrate. But the spermatozoa contain the enzyme, hyaluronidase, and this, if present in sufficient amount, quickly breaks up the jelly-like complex of hyaluronic acid by which the *discus* cells are glued together.‖

Nervous Mechanism of Coitus. Although, in both sexes, coitus is attended by a high degree of psychical excitement, yet its reflex part can be carried out when all impulses from the higher centres are cut off by section of the cord in the dorsal region. The centres presiding over the reflexes are situated in the lumbar and sacral spinal cord (ejaculation and erection respectively). Like the bladder, the external generative organs are supplied from two sets of nerve fibres—from the lumbar nerves through the sympathetic, and with parasympathetic fibres from the sacral nerves.

The fibres from the lumbar nerves arise, in the cat, from the second, third, and fourth, or the third, fourth, and fifth lumbar nerve roots, and, in the dog, from the thirteenth thoracic, and the first to the fourth lumbar roots. They run in the white rami communicantes to the sympathetic chain, whence they may take one of two paths.

(*a*) The great majority of the fibres run down the sympathetic chain to the sacral ganglia, whence fibres are given off in the grey rami communicantes to the sacral nerves ; their further course is by the pudic nerves.

(*b*) A few fibres go by the hypogastric nerves to the pelvic plexus.

---

* Bickers and Main. *J. Clin. Endocrin.*, 1941, **1**, 992.

† Parker. *Phil. Trans. R.S.*, 1931, **219B**, 381; Carleton and Florey. *J. Obstet. and Gynæc.*, 1931, **38**, No. 3.

‡ Walton. *Proc. R. S.*, 1927, **101B**, 303.

§ Hotchkiss *et al.* *Amer. J. Med. Sci.*, 1938, **196**, 362.

‖ McClean and Rowlands. *Nature*, 1942, **150**, 627.

Excitation of these fibres causes contraction of the arteries of the penis, and of the unstriated muscles of the tunica dartos of the scrotum. In animals which possess a retractor penis muscle, excitation of the lumbar nerves causes strong contraction of that muscle. The effect of stimulation of these nerves is, therefore, to inhibit erection.

The fibres from the sacral nerves can be divided into two classes—visceral and somatic. The visceral branches run in the pelvic nerves, or *nervi erigentes.* Stimulation of these nerves produces active dilatation of the arteries of the penis or vulva, and also inhibition of the unstriated muscle of the penis, of the retractor muscle of the penis when present, and of the vulva muscles. The somatic branches supply motor nerves to the ischio- and bulbo-cavernosi, as well as to the constrictor urethræ. In the female, they supply the analogous muscles, namely, the erector clitoridis (ischio-cavernosus) and the sphincter vaginæ (bulbo-cavernosus). Both sets of sacral fibres are therefore involved in the erection of the generative organs which accompany coitus.

The internal organs, *i.e.* the uterus and vagina in the female, and vasa deferentia, seminal vesicles, and uterus masculinus in the male, differ from the external organs in receiving no efferent nerve fibres from the sacral nerves.* They are supplied with fibres, which pass out through the ventral roots of the third, fourth, and fifth lumbar nerves (in the rabbit and cat), and run through the sympathetic to the inferior mesenteric ganglia, whence they proceed by the hypogastric nerves. On stimulating these fibres, two effects are produced on the uterus and vagina, namely, a contraction of the small arteries, leading to pallor of the organs, and a strong contraction of the muscular coats. Under some circumstances, stimulation of the sympathetic nerves may cause *relaxation* of the uterus, as also may the administration of adrenaline. In the vagina, the contraction can usually be seen to start from one end and spread to the other. The whole then remains for a time in a state of powerful tonic contraction, which affects both longitudinal and circular muscles. There is also secretion of Bartholin's glands.† In the male, stimulation of these nerves excites contraction of the whole musculature of the vasa deferentia and seminal vesicals which may be strong enough to cause emission of semen from the penis. These effects on the uterus and seminal vesicles are not abolished by injection of atropine.

The ovary in monkeys has a sympathetic nerve supply from the tenth thoracic to seventh lumbar levels of the spinal cord. ‡

Removal of the sympathetic system causes impotence in males if it causes ejaculation to fail, but in some human cases this result does not ensue§ ; it has no effect on the œstrous cycle, course of pregnancy or parturition in females,‖ though in some of these lactation is inhibited.

The course of the sensory fibres which represent the afferent side of the arc, from the generative organs to the lumbo-sacral cord, has not yet been fully made out, but it seems probable that it corresponds to the course taken by the efferent fibres.

An accessory genital muscle, the *retractor penis*, is found in the dog, cat, horse, donkey, hedgehog (not in the rabbit or man). It consists of a thin band of longitudinally arranged unstriated muscle, which is inserted at the attachment of the prepuce, and is continued backwards to the bulb, where it divides into two slips which pass on either side of the anus. A few striated fibres are found in the posterior part.

The *motor* fibres to the muscle are derived from the lumbar sympathetic (*i.e.* the upper set of nerve roots), and run to the muscle in the internal pudic nerve. The pelvic nerves, on the other hand, carry *inhibitory* impulses to the muscle, thus enabling the concomitant vascular dilatation to take effect in producing erection of the penis.

For Reference

Gruber (1933). The Autonomic Innervation of the Genito-urinary System. *Physiol. Rev.*, **13**, 497.

* Langley and Anderson. *J. Physiol.*, 1895, **19**, 122.

† Barrington. *Int. Mtschr. Anat. Physiol.*, 1913, **30**, 1.

‡ Zuckerman. *Anat. Rec.*, 1934, **58**, Sup. 43.

§ Takats and Helfrich. *J. Amer. Med. Ass.*, 1941, **117**, 20.

‖ Cannon *et al.* *Amer. J. Phys.*, 1929, **89**, 83 ; Bacq. *Amer. J. Phys.*, 1932, **99**, 444.

PREGNANCY

Fertilisation of the ovum takes place in the Fallopian tube. In the human subject, fertility is probably highest about ten days after the end of menstruation and lowest about a week before its onset. The data published by Siegel,[*] in which the leave periods of German soldiers during the 1914 war were correlated with the incidence of pregnancy in their wives, generally supports this view. More recently, deliberate records of dates of menstruation and of acts of intercourse place the high fertility period at the tenth to thirteenth days.[†]

Directly one spermatozoon has penetrated into the ovum, a membrane is formed round the yolk, which prevents the entrance of any other spermatozoa. The fusion of the male and female pronuclei is followed immediately by division of the fertilised ovum, so that, by the time it arrives in the uterus (about eight days after fertilisation), it consists of a mass of cells known as the morula. At this time, the ovum has a diameter of about 0·2 mm. Pregnancy in the human being lasts about 280 days, *i.e.* ten periods, after the last menstrual period.

With the arrival of the fertilised ovum in the uterus, extensive changes begin in the organs of generation. The virgin uterus cavity is about 2–5 c.c.; at term it is about 5,000–7,000 c.c., and the walls are thickened in proportion. In the hypertrophy, all its elements are involved, but especially the muscle cells. There may be formation of new muscle fibres, and each fibre also grows greatly in size during the first half of pregnancy (Fig. 678).

As the contents of the uterus increase in bulk the muscle fibres undergo an extension in length to accommodate the contents without any increase of intrauterine pressure, and continue always to keep a steady but gentle grip on the contents.

Anatomically the non-pregnant uterus is divisible[‡] into (*a*) the upper segment or *corpus* (divided in some animals into two horns), in which the embryo is implanted; (*b*) a physiological "retraction ring" or anatomical internal os, (*c*) a lower uterine segment or *isthmus*, and (*d*) a cervical part, in some animals containing a sphincter, and including the *external os*.

In the early stages of pregnancy the products of conception occupy the upper segment, but as growth proceeds, the products extend into the lower segment.

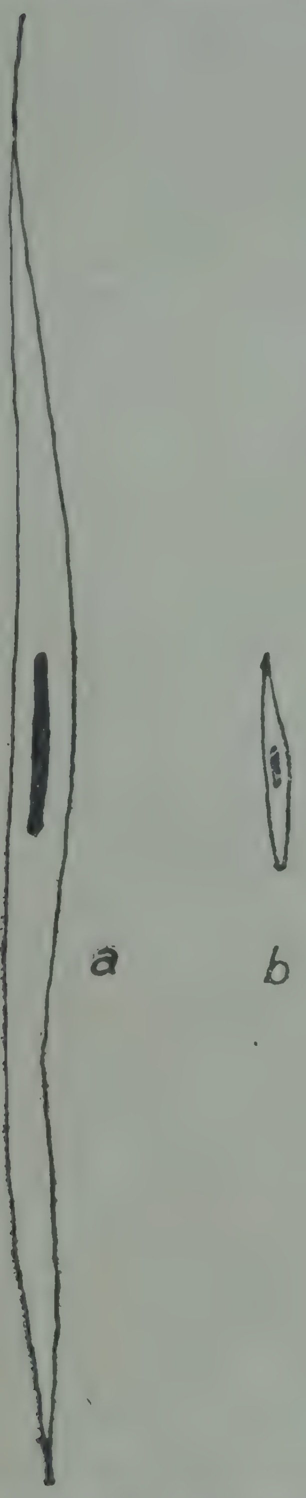


FIG. 678. Isolated Muscle Cells from the Uterus, showing the Hypertrophy during Pregnancy.

a. Fibre from uterus in ninth month of pregnancy. *b.* Fibre from a non-gravid uterus.

(After BUMM.)

During pregnancy there is a great growth of the blood vessels, which have to supply the growing wall of the uterus, and also, by means of the placenta, the nutritional needs of the developing fœtus. The increase chiefly affects the veins, which thus become capable of holding a considerable store of blood; the enlargement begins early in pregnancy, and while the fœtuses are still very small, continues more slowly while the fœtus grow rapidly, and finally, just before delivery, again increases quickly.[§] The rate of blood flow through the uterus increases at about the same rate as does the uterine blood content.[||]

---

[*] SIEGEL. *Münch. med. Woch.*, 1916, **63**, 748.

[†] PRYDE. *B.M.J.*, 1941 (i), 12.

[‡] DANFORTH, GRAHAM and IVY. *Q. Bull. Northwestern Med. Sch.*, 1941, **15**, 21.

[§] J. BARCROFT and ROTHSCHILD. *J. Physiol.*, 1932, **76**, 447.

[||] J. BARCROFT, HERKEL and S. HILL. *J. Physiol.*, 1933, **77**, 194.

Changes in the Uterine Mucous Membrane. After conception, the uterine mucous membrane shows a continuation of those hypertropic changes which we have seen to occur regularly during the premenstrual phase (Fig. 679). Within fourteen days, it has attained a thickness of ½ cm., and by the end of the second month ¾ cm. On section, it shows a compact layer, lining the cavity of the uterus, and outside this, abutting on the muscular tissue, is a spongy layer three times as thick as the compact layer. The superficial epithelium becomes flattened, loses its cilia, and degenerates. In the spongy layer, the uterine glands proliferate, the stroma cells are enlarged, and the blood capillaries are widely dilated. The stroma cells become converted into the large *decidual* cells. By the time the fertilised ovum arrives in the uterus, the process of hypertrophy of the layers of the mucous membrane has already made some progress. As it lies on the mucous membrane, the outermost cells of the developing ovum exercise a destructive influence on the adjacent cells of the mucosa, apparently through some sort of digestion, so that the ovum sinks into the membrane and reaches the sub-epithelial connective tissue. Round the margins of the depression which the ovum has made for itself, the mucous membrane grows over the protruding part of the ovum (Fig. 680). When this has taken place, the different parts of the mucous membrane receive different names. Since (in man) they are all to be cast off with the fœtus at birth, each part is spoken of as a decidua, that lining the main body of the uterus being known as the *decidua vera*, that covering the protruding part of the egg as the *decidua reflexa*, while that to which the egg is immediately attached is the *decidua serotina* or *basalis*. It is from the latter that the placenta is formed. By the end of the second week, the blood vessels in this situation are considerably enlarged. This enlargement proceeds, affecting especially the capillaries and veins, until these form venous sinuses at the junction between the mucous membrane and the muscular coat. Changes take place at the same time in the embryo. When it sinks into the mucous membrane it has a diameter of 1 mm. The blastoderm is fully formed, with its three layers; the yolk sac, the body cavity and the amnion are present. The outermost layer of the

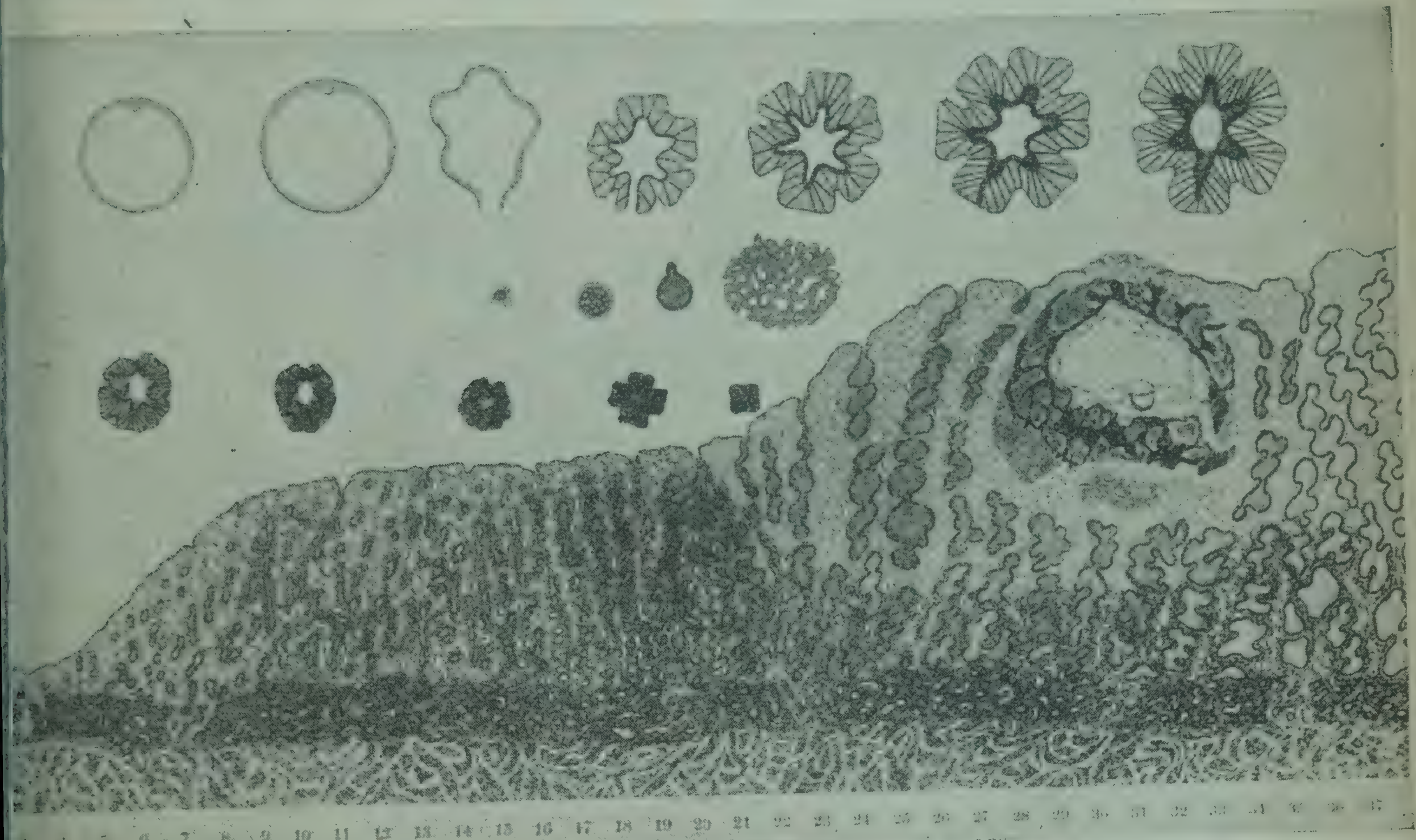


FIG. 679. Diagram showing Changes in the Human Endometrium and Ovary during first 37 days of Pregnancy. (After SCHRÖDER.)

epiblast becomes specially modified to serve for the nutrition of the embryo, and gives rise to numerous villi, the chorionic villi, so that the whole ovum has a shaggy appearance. Since this tissue takes no part in the further development of the embryo, but serves simply for its nutrition, it is often spoken of as the *trophoblast*. With the formation of fœtal blood vessels, these penetrate into the villi, together with mesoblast. The villi grow into the venous spaces, especially in the basal part of the decidua, so that, at this period, the fœtal villi are immersed in maternal blood, the fœtal blood vessels being separated from the maternal blood by a double layer of epithelium, one layer of which is maternal and the other fœtal in origin. Later, these cells become reduced to a single layer.

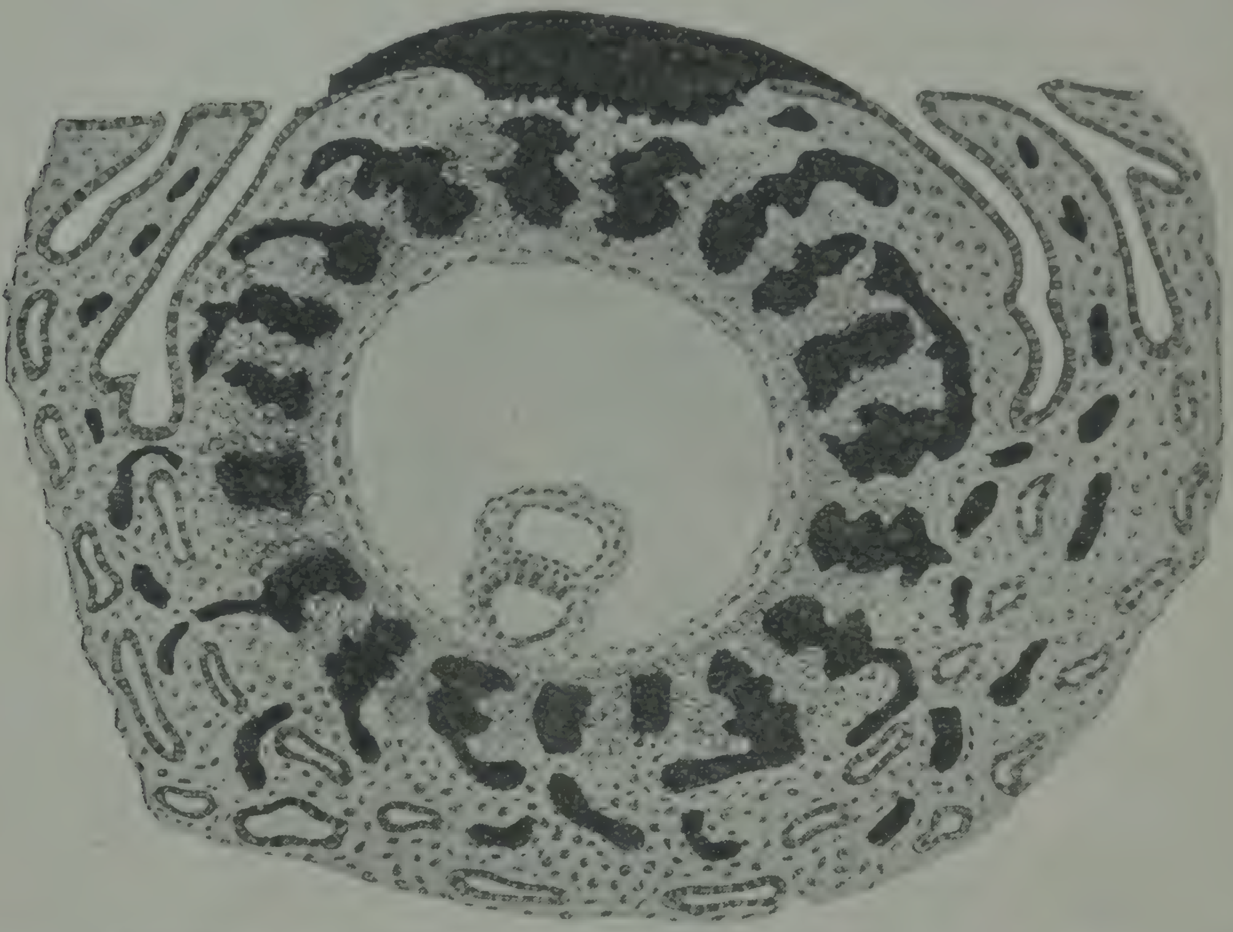


FIG. 680. Diagram to illustrate the Imbedding of the Ovum in the Decidua, and the First Formation of the Fœtal Villi in the form of a Syncytial *Trophoblast* (derived from the Outer Layer of the Ovum) which is invading Sinus-like Blood Spaces in the Decidua. (After T. H. BRYCE.)

Hormone Factors. Pregnancy is accompanied by pronounced changes in the amount and distribution of various sex hormones. Even in the early stages there is an excretion in the urine of a gonadotropic hormone similar to, but not necessarily identical with, the luteinising hormone (I.C.S.H.) of the anterior pituitary. It is called *chorionic gonadotropin* (P.U.). It reaches its maximum at about the eighth week, falls till the tenth or eleventh week to a lower level, then stays constant till delivery, and disappears a few days later.

The presence of a substance having gonadotropic actions in the urine of pregnant women,* has led to a test for pregnancy.† To 1 to 3 c.c. of urine, previously extracted with ether, glucose is added; it is then injected twice daily in various doses into immature female mice. Within 72 hours, at least one out of five mice shows the positive reaction—viz. one or more corpora lutea or blood points (hæmorrhagic follicles). An even more reliable test is the Friedman test, which utilises adult female rabbits, segregated for a month. Two i–v. injections of about 8 c.c. of morning urine are made on two days, and on the third day the animal is killed and the ovaries examined for hæmorrhagic follicles or follicular craters. If the urine is injected into hypophysectomised rats the action is only luteinising; the follicular maturation seen in immature animals, being secondary to an effect on their own pituitaries. The P.U. principle is thus mainly luteinising in its action. It is assayed by studying vaginal cornification. The serum of pregnant mares (P.M.S.) contains both follicle-stimulating and luteinising hormones; the former can be assayed by finding the increase produced in the weight of the ovaries of immature animals.

---

* ASCHHEIM and ZONDEK. *Klin. Woch.*, 1928, **7**, 1404.

† ZONDEK. *Klin. Woch.*, 1931, **10**, 1484.

A third gonadotropic principle (C.U.) is found in the urine of ovariectomised women, and after the menopause. It is mainly follicle-stimulating and may be a true pituitary hormone.*

International standards containing arbitrary fixed amounts of active principle have been laid down for the gonadotropic luteinising substance of human urine of pregnancy ("chorionic gonadotrophin") (0·1 mg.) and for the (follicle-stimulating) gonadotropic substance of pregnant mare's serum (0·25 mg.). These standards are kept by the Medical Research Council.†

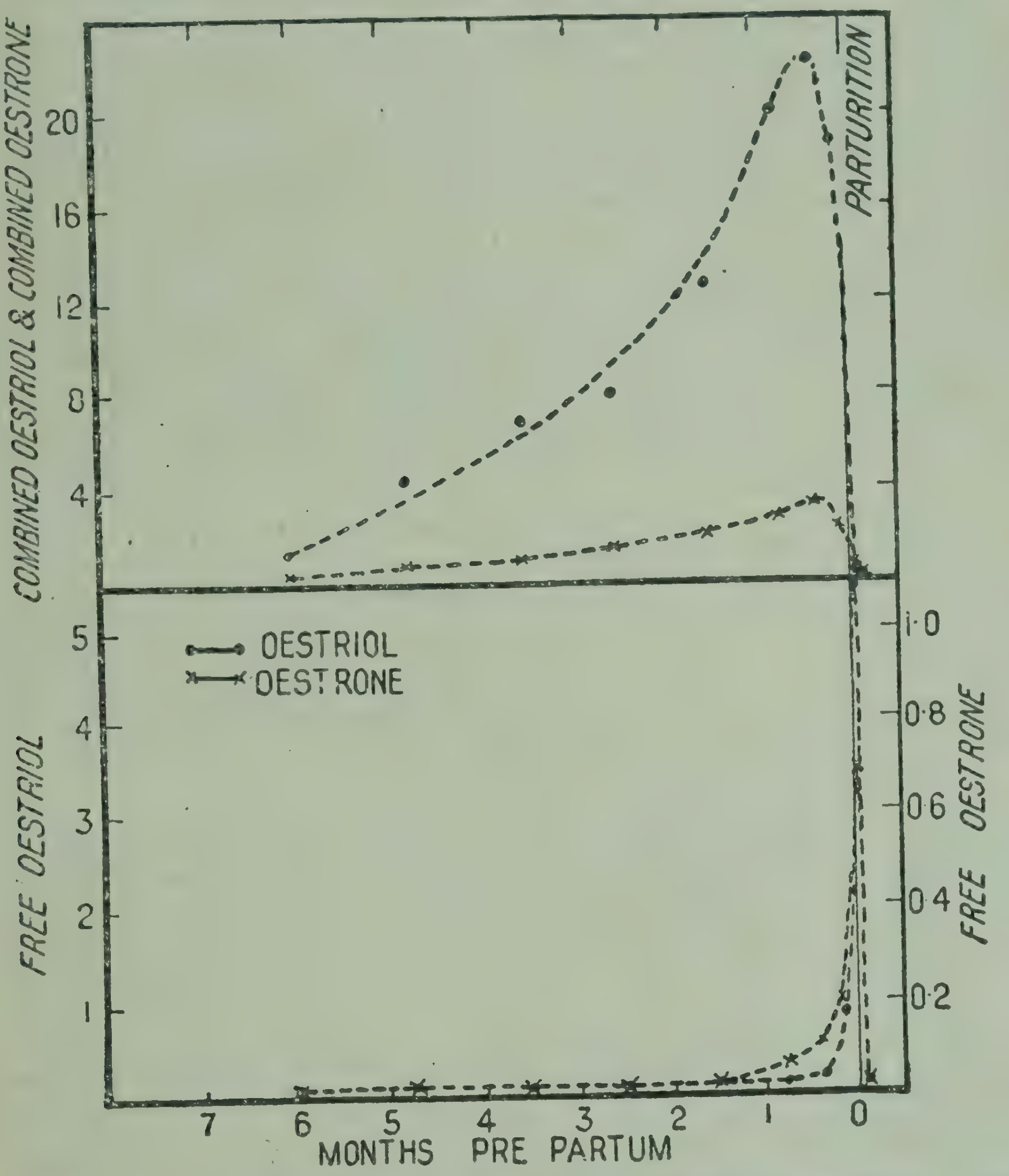


FIG. 681. Excretion of Œstrogen in the Urine during Pregnancy. Œstriol and Œstrone in free and combined forms are expressed as mg. per 24 hours. (COHEN, MARRIAN and WATSON. *Lancet*, 1935 (i.), 674.)

Another test for pregnancy ‡ consists in giving to the patient three intramuscular injections of 1 mg. of œstrone in oil over a period of five days. If menstruation does not begin within 24 hours after the last injection, the patient is considered to be pregnant.

After the eighth week, œstrogen in combined form (as a glucuronide) is excreted in increasing amount right up to the end of term.§ The enlargement of the corpora lutea, and the increase in excretion of pregnandiol glucuronidate in the urine from the eighth week onwards, show an increasing formation of progesterone. The combined action of the œstrogens and progesterone brings about enlargement and differentiation in structure of the mammary glands, and uterus, and other changes already referred to. Œstrogens can in large doses suppress lactation, but in small doses encourage growth of the mammary duct system and may sometimes stimulate the pituitary anterior lobe to produce prolactin.* As soon as labour sets in there is a sudden drop in the excretion of the combined and relatively inactive urinary œstriol and œstrone, and a sharp rise in the free and active forms (Fig. 681).†

---

* BROWNE. *Can. J. Res.*, 1933, **8**, 180.

† *Bull. Health Organisation, League of Nations*, 1939, **8**, 884, 898.

‡ GARRETT, *Amer. J. Surg.*, 1948, **76**, 261.

§ COHEN and MARRIAN. *Bioch. J.*, 1936, **30**, 57.

FUNCTIONS OF THE PLACENTA. Besides playing an important part in the nourishment of the fœtus, it now seems almost certain that the placenta is a site for production of all the important hormones of pregnancy. The P.U. factor is almost certainly produced by chorionic tissue, since it is also excreted in such conditions as chorionepithelioma, in which there is no fœtus. In some species, including the human, pregnancy is not terminated by oöphorectomy, after which œstrogens and pregnandiol glucuronide continue to be excreted. Lastly, Newton has shown,‡ in the mouse, that if the fœtuses are destroyed, but the placentæ remain, œstrus is suppressed, the mammary glands develop (independently of the ovaries), and finally the placentæ are delivered at the normal term.

Nutrition of the Embryo. At the earliest period of its development, the fertilised ovum is dependent for its nourishment on the remains of the cells of the discus proligerus adhering to it, or on the fluid of the Fallopian tube in which it is immersed. The first blood vessels which are formed serve to take up nourishment from the yolk sac. In man, this source of supply is insignificant, and, from the second week onwards, blood vessels traversing the chorionic villi come into close relation with the maternal blood, from which, henceforth, the whole growth of the fœtus is to be maintained by a special development of these connections in the placenta.

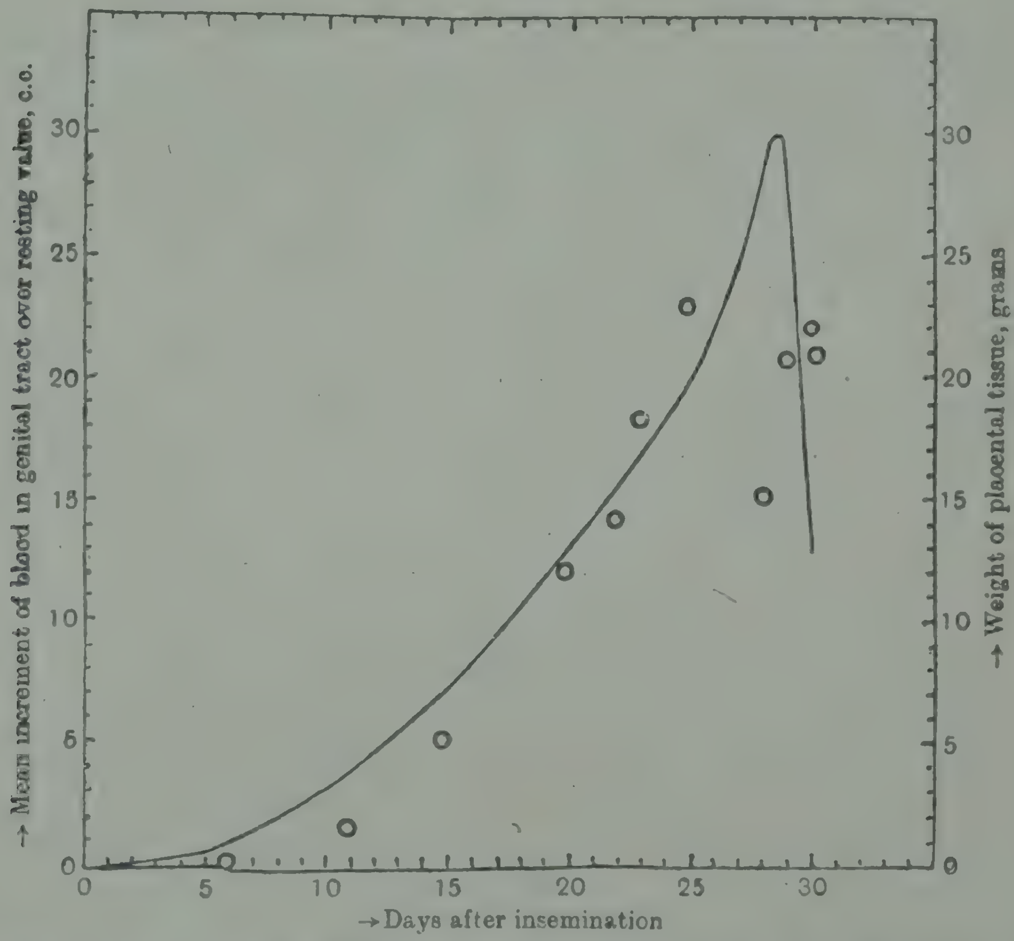


Fig. 682. Relation between Placental Weight (circles) and Volume of Blood in Uterus (line) in pregnant rabbit. (After Barcroft and Rothschild.)

In the fully formed fœtus, blood passes from the fœtus to the placenta by the umbilical artery, and is returned by the umbilical veins. There is no direct communication between fœtal and maternal circulations. The placenta represents the fœtal organ for respiration, nutrition, and excretion. Thus, the umbilical artery carries to the placenta a dark, venous blood, which in this organ loses carbon dioxide and takes up oxygen, so that the blood of the umbilical veins is arterial in colour. The oxygen requirement of the fœtus is small, probably less than half of that of the same weight of maternal tissue. It is protected from all loss of heat, movements are sluggish, or for the most part absent, and the only oxidative processes are those required in the building up of the developing tissues. On the other hand, the fœtus has need of a rich supply of foodstuffs, which it must obtain through the placental circulation. It is imagined that the epithelium covering the villi represents a membrane through which the necessary foodstuffs and oxygen pass from the maternal to the fœtal blood, and carbon dioxide, urea and other metabolic products in the reverse direction. The placenta contains a considerable amount of glycogen and fat, which may perhaps act as stores of material for the use of the fœtus in emergency.*

---

* Folley. *B.M.J.*, 1941 (i), 40.

† Cohen, Marrian and Watson. *Lancet*, 1935 (i.), 674.

‡ Newton. *J. Physiol.*, 1935, **84,** 196 ; *Anat. Rec.*, 1938, **72,** 333.

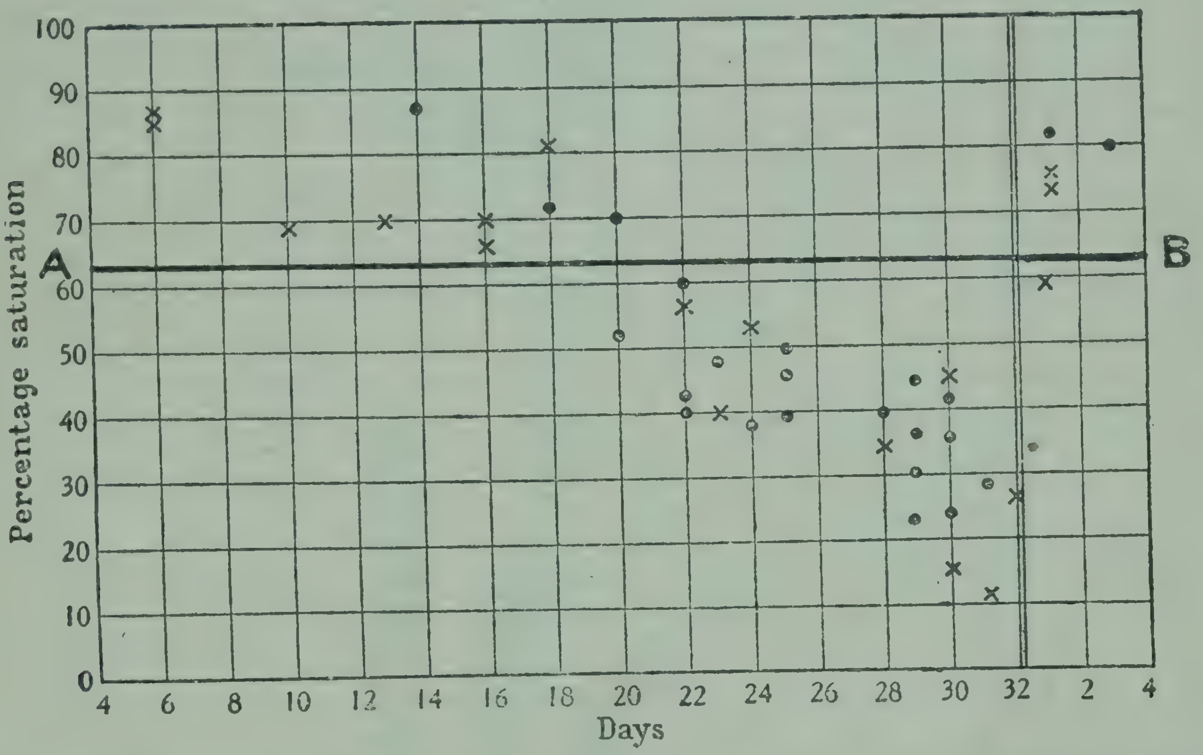


FIG. 683. Percentage Saturation with Oxygen of Maternal Blood leaving the Pregnant Horn of the Uterus at different stages of Pregnancy in the Rabbit. AB represents the average saturation of blood leaving the non-pregnant horn. Parturition on thirty-second day. (BARCROFT, FLEXNER, HERKEL, MCCARTHY and MCCLURKIN.)

As might be anticipated from their rapid rate of growth, the amount of blood in the uterus, placenta and fœtus rapidly increases during pregnancy; the fœtal goat, for instance, shows an increase of blood volume from about 16 c.c. at the sixty-eighth day to 300 or 400 c.c. at term. Expressed as a percentage of the weight of the fœtus, however, the amount falls slightly during pregnancy from about 25 to about 14 per cent., but if expressed as a percentage of the weight of fœtus plus placenta, the amount is more nearly constant at about 9 per cent. In other words, as pregnancy advances, the fœtus grows more rapidly than the placenta, but the blood volume roughly keeps pace with the growth of the combined uterine contents.

Two factors are clearly of importance in facilitating the exchange between fœtal and maternal bloods, viz. the circulation rate in the fœtus, and the circulation rate in the maternal placenta. In the fœtal sheep and goat, the output of the fœtal heart provides a fairly constant volume of blood per unit weight of fœtal tissue during the latter half of pregnancy, but the volume per unit weight of whole tissue to be circulated, *i.e.* of fœtus plus placenta, steadily increases during the latter half of pregnancy, the volume per gramme per minute being about three times as much at full term as it was at about half term.

---

* ELLIOTT, HALL and HUGGETT. *J. Physiol.*, 1934, **82**, 160.

On the maternal side there is a great increase in the vascularity of the uterus, which, in fact, runs closely parallel to the increase in weight of the placentæ (Fig. 682). Since the growth of the fœtus soon outstrips that of the placenta, however, it would appear that the maternal blood circulation rate tends to lag behind the fœtal one. The relation between these two circulations becomes of paramount importance in connection with the supply of oxygen to the fœtus since, as pregnancy advances, the maternal venous blood leaving the placenta becomes less and less saturated with oxygen (Fig. 683).

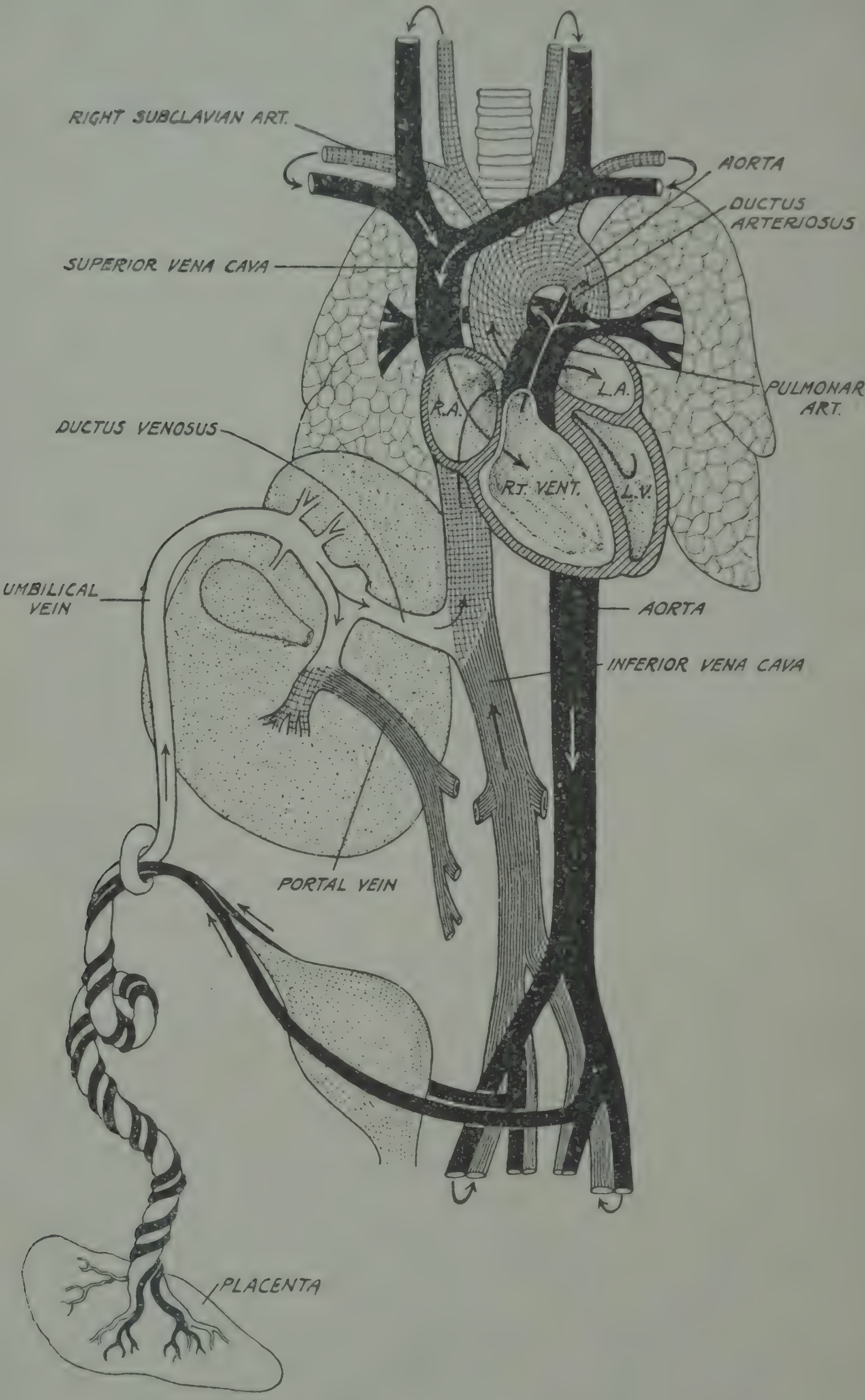

RIGHT SUBCLAVIAN ART.
AORTA
DUCTUS ARTERIOSUS
SUPERIOR VENA CAVA
PULMONARY ART.
L.A.
R.A.
DUCTUS VENOSUS
RT. VENT.
L.V.
UMBILICAL VEIN
AORTA
INFERIOR VENA CAVA
PORTAL VEIN
PLACENTA

Fig. 684. Diagram of the Fœtal Circulation. (After Gray.)

The Fœtal Circulation. In the fœtus, from the middle of intra-uterine life, we find certain arrangements of the circulation which have long * been supposed to be directed to providing the heart itself, and the forepart of the body, especially the rapidly growing brain, with oxygenated blood, while the less important tissues of the limbs and trunk receive almost venous blood (Fig. 684). These old beliefs have been vindicated by direct experiment, in which radio-opaque solutions have been injected into the fœtal circulation, and X-ray cinematograph records taken.† As the arterial blood of the umbilical vein passes from the placenta, *viâ* the *ductus venosus*, into the inferior cava, it is mixed with the sluggish stream of venous blood, returning from the lower limbs and lower part of the trunk. Some of the umbilical vein blood, of course, can pass directly into the liver. The greater part of it, however, traverses the ductus venosus, to enter the inferior vena cava, by which it is carried to the right auricle. Here the bulk of it impinges on the Eustachian valve, and is directed thereby through the *foramen ovale* into the left auricle, whence it passes into the left ventricle, to be driven into the aorta. By the aorta, this partially arterialised blood is carried chiefly to the coronary vessels of the heart, and to the head and fore limbs. A small fraction of it, however, passes on through right auricle and ventricle to the pulmonary artery. The venous blood from the fore parts is also carried by the superior vena cava to the right auricle, right ventricle, and pulmonary artery. Of the but slightly arterialised mixture of bloods thus carried to the pulmonary artery, only a small part passes through the lungs, the greater part traversing the patent *ductus arteriosus* to be discharged into the descending aorta, whence about a third of it flows to the lower limbs and trunk, and two-thirds to the placenta by the umbilical arteries. In the fœtus, therefore, the work of the circulation is largely carried out by the right ventricle. The great thickness of the left ventricular walls, which is so characteristic of the adult, does not become evident until shortly before birth. The umbilical artery, accordingly, carries chiefly venous blood to the placenta, where it is arterialised by interchange with the maternal blood, and the arterialised blood is then returned to the fœtus by the umbilical vein—a state of affairs comparable with that resulting from passage of blood to and from the gills of fish. Conversely, of course, the maternal blood conveyed to the uterus and placenta is arterial, and that returning is venous.

J. Barcroft found the following oxygen contents of the various fœtal bloods in the sheep :—

Blood of umbilical vein	= 79	per cent.	oxygen	saturation
,, thoracic inf. ven. cava . . .	= 56	,,	,,	,,
,, carotid artery	= 41	,,	,,	,,
,, jugular vein	= 33	,,	,,	,,
,, aorta (umbilical artery) . .	= 23	,,	,,	,,

He believes that, as postulated, there is a crossing of the circulations, though not a complete one, and estimates that about two-thirds of the blood from the " upper " circulation passes on to the lower one, while of the blood returning from the lower circulation two-thirds pass on to the upper.

SUPPLY TO THE FŒTUS. The rate at which oxygen is acquired by the blood circulating through the fœtal side of the placenta will depend especially on the relative rates of the fœtal and maternal placental circulations, and on the oxygen tensions on the two sides of the membranes. The most favourable conditions would be when the circulation rates on both sides were rapid, and the difference between the oxygen tensions of maternal and fœtal bloods great. Such information as we have is derived chiefly from the study of the rabbit, sheep and goat, and it is uncertain how far the conclusions drawn from these species can be transferred to man.

We have seen that, as pregnancy advances, the circulation rate on the maternal side fails to augment at the same rate as that on the fœtal side, which increases even more rapidly than does the weight of tissue to be supplied. In view of this it is not surprising to find that the percentage oxygen saturation of the blood leaving the uterus (much of which comes from the placenta) steadily falls towards the end of pregnancy, when it may reach a very low figure, *e.g.* 20 per cent., as against about 65 per cent. saturation in the non-pregnant uterine venous blood, or 75 per cent. in the post-partum uterine venous blood.

---

* SABATIER. Traité complet d'anatomie, 1791, ii., 493.

† BARCLAY, BARCROFT, BARRON and FRANKLIN. *Brit. J. Radiol.*, 1939, **12**, 505.

The surprising fact which emerges from these results is that the fœtal blood coming from the placenta has so much greater a saturation with oxygen than has the maternal venous blood. Yet it has been shown that the gradient of tensions is such that the exchange can be accounted for by diffusion across the intervening membrane,* *i.e.* the *tension* in the fœtal blood returning by the umbilical vein is *not* higher than that in the maternal venous blood. The key to this apparent discrepancy has been found by the study of the dissociation curves of fœtal and maternal bloods. It has been shown † that the curve of the fœtal blood is steeper than that for maternal blood, *i.e.* at a given oxygen tension the fœtal blood takes up more oxygen than does the maternal blood (Fig. 685). This appears to be partly because the maternal blood, owing to raised H-ion concentration, has a lowered oxygen affinity, but chiefly because the hæmoglobin of fœtal blood is different from that of the adult, in having a higher oxygen affinity.

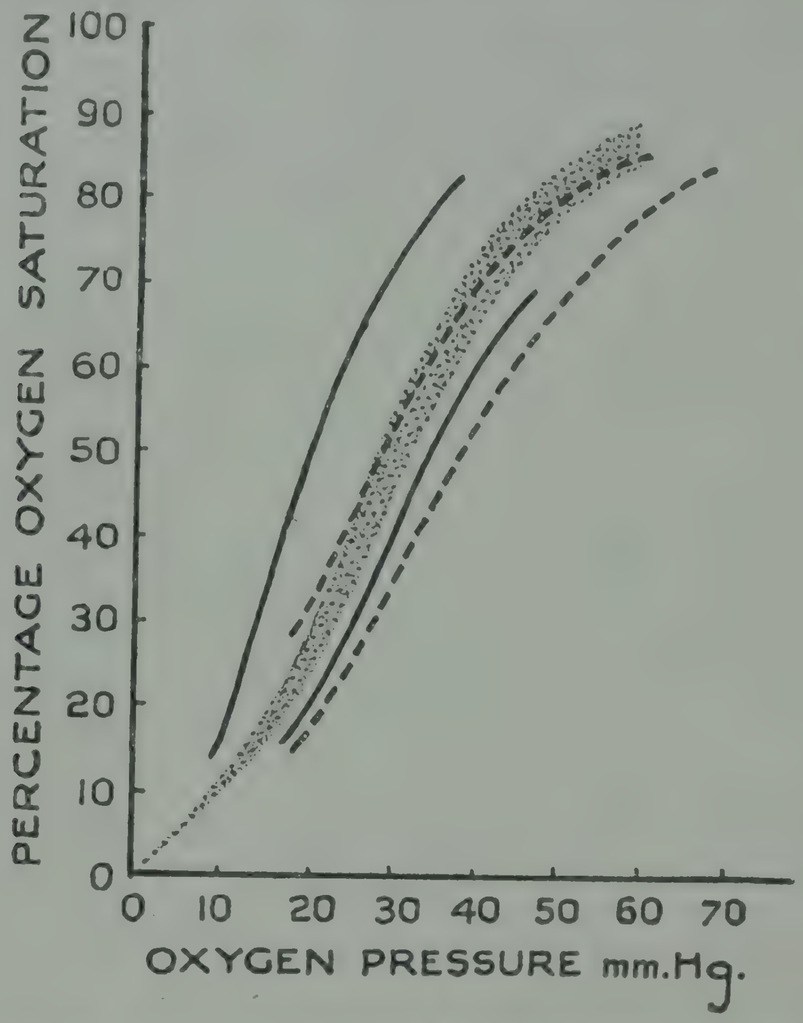


FIG. 685. Oxygen Dissociation Curves of Fœtal and Maternal Goats. The two left curves are from fœtal and the two right from maternal bloods. The shaded area covers the range of dissociation curves of normal adult goats. Stage of gestation: continuous lines, $19\frac{3}{7}$ weeks; interrupted lines, $14\frac{1}{7}$ weeks. (After BARCROFT and others. *Journ. Physiol.*, 1934, **83**, 192.)

Various relations are possible between the directions of flow in the maternal and fœtal vessels of the placenta, as suggested in Fig. 686. That which would give the highest oxygen tension in the fœtal blood would be as shown in A (the condition in the rabbit) in which the fœtal blood on leaving the capillary is in relation with the maternal blood just entering; in B, on the other hand, the fœtal blood would be in equilibrium with maternal venous blood. Probably in man the arrangement is as in C, which, supposing that diffusion equilibrium were reached, would result in a fœtal oxygen tension in the umbilical vein blood higher than that in the maternal venous blood, but lower than that in the maternal arterial blood.

---

* HUGGETT. *J. Physiol.*, 1927, **62**, 373.

† HASELHORST and STROMBERGER. *Ztsch. Geb. Gyn.*, 1930, **98**, 49; 1931, **100**, 48; BARCROFT *et al.* *J. Physiol.*, 1934, **83**, 192.

Actual determination of the oxygen saturation of the various bloods supports this conclusion. J. Barcroft found in the goat and sheep that the percentage saturations are somewhat as follows at about the middle of fœtal life :—

Arterial blood to uterus	95	per cent. oxygen saturation	Maternal circuit.
Venous blood from uterus	60	,, ,, ,,	
Umbilical artery (to placenta)	15–45	,, ,, ,,	Fœtal circuit.
Umbilical vein (from placenta)	70–80	,, ,, ,,	

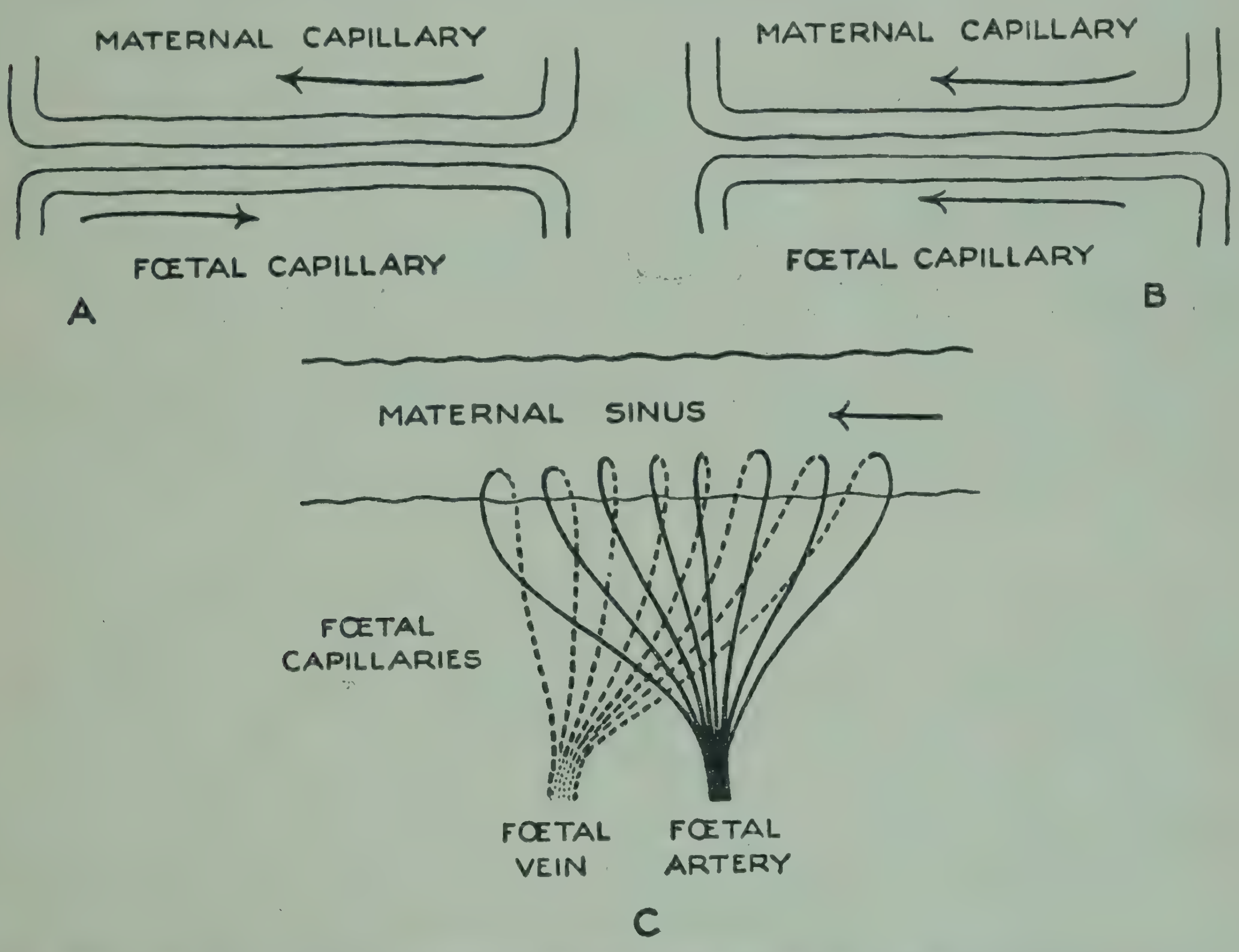


Fig. 686. Possible Relationships between Maternal and Fœtal Circulations. A, opposite flow ; B, parallel flow ; C, " perpendicular " flow. (Newton, " Recent Advances in Physiology.")

As regards the transfer of carbon dioxide, similar considerations will apply. There seems no reason to invoke any other aid than that of diffusion, for effecting the transfer from fœtal to maternal blood.

Changes at Birth. The first breath taken by the newborn child is due to the effect of slight, or sometimes definite, asphyxia consequent upon separation of the placenta from the uterus and to a lesser extent to the influence of cold and of drying of the skin. With sufficiently advanced fœtuses, opening of the uterus or of the uterus and sac in air leads to respiration, but opening under warm saline solution does not. Ligation of the umbilical cord or of the arteries supplying the uterus invariably causes respiration.* Respiratory movements, probably also asphyxial in origin, always occur late in intra-uterine life, but the amniotic fluid thus aspirated into the lungs is reabsorbed. The respiratory centre of the fœtus acts less readily than that of the adult animal.† This appears to be due to an inhibitory influence exerted by the forebrain.

As the time for birth approaches, the oxygen saturation of the blood in the umbilical vein in the human is about 50 per cent., and that in the fœtal carotid artery falls to a level not much above that of the umbilical artery, so that the respiratory centre is now more easily released from inhibition.*

---

* Corey. *J. Exp. Zool.*, 1932, **61**, 1.

† Huggett. *J. Physiol.*, 1930, **69**, 144.

With the first breath taken by the newborn child, several mechanical conditions of the circulation are modified. The resistance to the blood flow through the lungs being diminished, the blood passes from the pulmonary arteries through the lungs into the left auricle. The pressure in the left auricle is thus raised, while that in the right auricle falls, so that the foramen ovale is maintained closed. The oxygen saturation of the arterial blood of infants rises within an hour and continues to improve so that the adult level is reached in from one to two weeks.† Even before birth, proliferation of the lining membrane may be seen both in the ductus arteriosus and in the ductus venosus; and with the mechanical relief of the vessels afforded by respiration and the changed conditions of the fœtus, this proliferation goes on to complete obliteration of the vessels.

For Reference

J. Barcroft (1936). Fœtal Circulation and Respiration. *Physiol. Rev.*, **16**, 103.
J. Barcroft (1946). Prenatal Life. Vol. I. (Oxford: Blackwell.)
Barron (1944). The Changes in the Fetal Circulation at Birth. *Physiol. Rev.*, **24**, 277.
Harding (1925). Metabolism in Pregnancy. *Physiol. Rev.*, **5**, 279.
Huggett (1941). The Nutrition of the Fetus. *Physiol. Rev.*, **21**, 438.
Needham (1925). The Metabolism of the Developing Egg. *Physiol. Rev.*, **5**, 1.
Needham (1931). Chemical Embryology. (Cambridge Univ. Press.)
Nelson (1936). Endocrine Control of the Mammary Gland. *Physiol. Rev.*, **16**, 488.
Newton (1938). Hormones and the Placenta. *Physiol. Rev.*, **18**, 419.
Snyder (1938). Factors Concerned in the Duration of Pregnancy. *Physiol. Rev.*, **18**, 578.

PARTURITION

In the first half of pregnancy the uterus is quiescent, and this is attributed to the inhibitory effect on it of progesterone. It is also unresponsive to oxytocin. Later, contractions of rhythmic character appear, at first feebly, then more powerfully. These contractions ordinarily give rise to no sensations, but are felt when they are augmented in consequence of reflex stimulation. During the greater part of pregnancy they have little or no effect on the contents of the uterus. During the last weeks or days of pregnancy, however, these contractions, which have now become more powerful, have a distinct physiological effect. Not only do they, by pressing on the fœtus, cause it, in most instances, to assume a suitable position for its subsequent expulsion, but, affecting the whole body of the uterus, including the longitudinal muscular fibres surrounding its neck, they assist the general enlargement of the organ in dilating the internal *os uteri*, so that the upper part of the cervix is obliterated and drawn up into the body of the uterus some little time before labour has commenced. Towards the end of term also, the uterus becomes sensitive to oxytocin, and this may be a fact of importance.

With these changes in the uterus are associated changes in the round ligaments and in the vagina and vulva. The muscular fibres of the round ligaments become much hypertrophied and lengthened, and these structures can therefore appreciably aid the uterine contractions in the subsequent expulsion of the fœtus. The vaginal walls become of looser texture, and so afford less resistance to distension during the passage of the fœtal head.

Considerable discussion has taken place as to the cause for the onset of labour or parturition, which in man occurs at a nearly constant period of 281 days after conception. It is perhaps significant that the period of gestation is a multiple of the sex cycle (œstrous in animals, menstrual in women). Most of the explanations which have been suggested, such as the great irritability of the uterus at the termination of pregnancy, the loosening of the fœtal membranes, the return of the menstrual congestion after ten missed periods, thrombosis of the placental sinuses, simply replace one question by another. Doubtless the reason why labour occurs at a definite time must be sought for in extrinsic, rather than in uterine, changes. It is probably hormonal in causation, and the problem may be viewed from an interesting angle if we inquire why it is that the presence of the fœtus is tolerated for so long. The quiescence of the uterus is now generally attributed to the production of progesterone; if, in the rabbit, ovulation is artificially induced on the twenty-fifth day of pregnancy, by injection of a dose of the luteinising principle of human pregnancy urine, it is found that parturition does not occur at the normal time of thirty-two days, but at the fortieth day. This is associated with the presence of maximal-sized corpora lutea at the thirty-second day, whereas in a normal pregnancy these are showing degeneration at that time.

---

* Barcroft, Barron, Cowie and Forsham. *J. Physiol.*, 1939, **96**, 12 P.

† Smith and Karplan. *Amer. J. Dis. Childr.*, 1942, **64**, 843.

The presence of the fœtus is not a necessary condition of the continuance of pregnancy, but the placenta appears to be important. If in pregnant mice the fœtuses are destroyed at the twelfth to fifteenth day of pregnancy, the placenta being retained, "pseudo-parturition" with delivery of the placentæ takes place at normal full term,* and is followed, as is normal delivery, by an œstrus period. This indicates that the placenta plays an important part in inhibiting œstrus and so in determining the date of parturition: it also suggests that an œstrus-producing factor is operative at about that time. It is attractive to speculate that the sudden drop in the excretion of combined œstrin and the accompanying increase in the excretion of free œstrin which immediately precede delivery in the human subject may be decisive factors, since it has been shown that the injection of large doses of œstrin into pregnant animals at any stage leads to abortion.

One view at present is that an important factor in the onset of parturition is a reduction in the inhibition exercised up to that time by the presence of progesterone, and so an increasing sensitiveness of the uterus to the action of oxytocin from the posterior pituitary. It is unquestionable that the administration of oxytocin causes rhythmic contractions of the parturient uterus very similar to those of normal labour (Burn and Bourne †), and it is also true that in some species abortion follows removal of the ovaries or of the corpora lutea. There is also some evidence that a liberation of oxytocin from the posterior pituitary may be a decisive factor. Stimulation of the supraoptico-hypophyseal tract results in increased uterine contractions, and a reflex secretion of oxytocin has been described as a result of distension of the genital canal in rabbits.‡ Removal of the pituitary, however, does not always affect the time of onset of parturition § though there is evidence of difficulty in the deliveries. Moreover, although it has been shown that oxytocin produces greater uterine responses when œstrin is also present,‖ this has not been substantiated by experiments on excised uteri treated with these substances.¶

According to Moir, the uterine contraction is a peristaltic one, passing downwards towards the cervix, the contraction of which comes last, and is feeble. The cervix responds differently from the body of the uterus to drugs, and possibly has a different innervation. It has been shown by Newton ** that the cornua of the pregnant goat's uterus respond to oxytocin whereas the cervix is insensitive, while adrenaline relaxes the excised cornua and contracts the cervix.

---

[*] Newton. *J. Physiol.*, 1935, **85,** 196.

[†] Burn and Bourne. *J. Obstet. and Gyn.*, 1927, **34,** 249; *B. M. J.*, 1930 (ii.), 87.

[‡] Ferguson. *Surg., Gynæc. Obstet.*, 1941, **73,** 359; Harris. *B.M.J.*, 1948, February 21st, p. 339.

[§] Dott. *Q. J. Exp. Phys.*, 1923, **13,** 241; Smith, *Am. J. Phys.*, 1932, **99,** 345; Allan and Wiles, *J. Physiol.*, 1932, **75,** 23.

[‖] Marrian and Newton. *J. Physiol.*, 1935, **84,** 133.

[¶] Newton. *J. Physiol.*, **79,** 301.

[**] Newton. *J. Physiol.*, 1934, **81,** 277.

To understand the physiology of labour it must be realised that not only do waves of contraction pass over the uterine muscle, to be followed by relaxation, but also that adjustments in the length of the muscle fibres occur, by which a "grip" is kept on the contents, should these change in bulk. In other words, "slack" is taken up or paid out according to conditions. Actual parturition, in the woman, is generally divided into three stages. In the first stage the contractions chiefly act in dilating the os uteri. The bulk of the expulsive work is done by the upper segment. A wave of contraction starts at the upper part of the corpus, passes to the lower segment and on to the cervical portion.* The head is pushed down against the latter, and so at the passing off of the contraction, some slack is taken up by the upper segments and some paid out by the cervix, so that the head, once down, cannot return: there is, in the second place, a passive dilatation by the pressure of the fœtal bag of membranes, filled with amniotic fluid, which is forced down as a fluid wedge into the os by the contractions of the uterine corpus. The uterine contractions are essentially rhythmical, being feeble at first, and increasing gradually in intensity to a maximum which endures a certain time, and then gradually subsides. The frequency and duration of the contractions increase as labour advances.

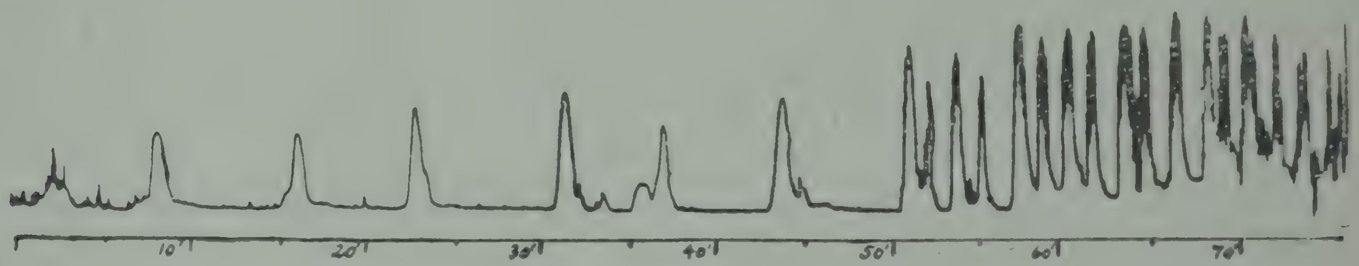


FIG. 687. Record of intra-uterine pressure of woman in labour. Note first stage pains up to the 50′ mark. Second stage pains recorded during next twenty-five minutes; they begin suddenly. The head of the child was born at 75′ where the pressure momentarily drops, and the body just after the last rise of the lever. The patient was a multipara. (BOURNE and BURN.)

Finally the cervix and os become fully dilated. The membranes generally rupture towards the end of the first stage of parturition.

The second stage begins as soon as the os uteri is fully dilated and the fœtal head has entered the pelvis, when the contractions change in character, being much more prolonged and frequent ("labour pains") and attended by more or less voluntary contractions of the abdominal muscles. This action of the abdominal muscles is associated with fixation of the diaphragm and closure of the glottis, so that pressure is brought to bear on the whole contents of the abdomen, including the uterus. After each wave has pressed the head down a little further, a relaxation follows: in this the head ascends a little, but not much, because of the progressive adjustment in length of the uterine fibres. No expelling force can be ascribed to the vagina, since it is too greatly stretched by the advancing fœtus. In this way, the fœtus is gradually thrust through the pelvic canal, dilating the soft parts which impede its progress, and is finally expelled through the vulva, the head normally being born first (Fig. 687).

A third stage of labour is generally described. This consists first in a considerable closing-in of the uterus, until it can regain grip of the greatly reduced contents; then, about twenty to thirty minutes after the birth of the child, a renewal of uterine contractions results in the expulsion of the placenta and decidual membranes, and the closing, by compression, of the venous sinuses.

Nervous Mechanism. We possess little experimental knowledge of the nervous mechanism of parturition. Goltz observed the normal occurrence of heat, impregnation, and parturition in a bitch whose spinal cord had been completely divided in the dorsal region during the previous year. On the other hand, destruction of the lumbo-sacral cord completely abolishes the normal uterine contractions of parturition, so that this act must be regarded as essentially reflex, presided over by a controlling " centre " in the grey matter of the cord. The activity of the centre can be inhibited or augmented by impulses arriving at it from the peripheral parts of the body, as by the stimulation of sensory nerves, or from the brain, as under the influence of emotions. The nerve paths from the centre to the uterus have been already described.

---

* IVY. *Amer. J. Obstet. and Gynec.*, 1942, **44**, 952.

THE PUERPERIUM. After delivery the uterus undergoes involution, and lactation soon commences. In the latter phenomenon the lactogenic hormone of the pituitary plays a decisive *rôle*. Suckling of the child is important in accelerating the involution of the uterus, which is caused to enter into powerful rhythmic contractions each time the child is put to the breast (Fig. 688). Excretion of œstrin and of gonadotropic hormones in the urine rapidly falls off and reaches the normal level in a week or ten days. Menstruation recommences when lactation is finished.

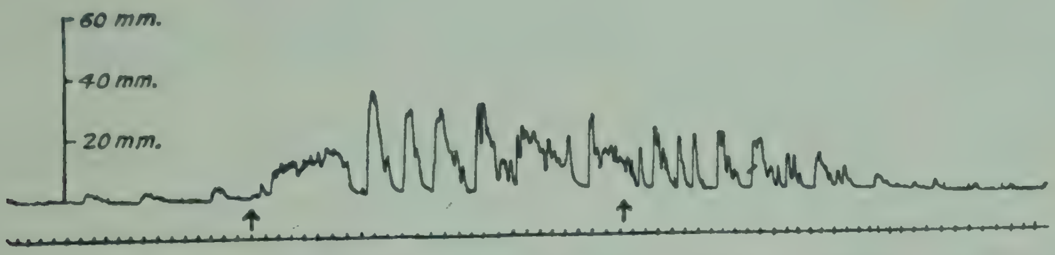


FIG. 688. Contractions of Human Puerperal Uterus. Suckling (between the arrows) caused increased contractions. (After MOIR. *Edin. Med. J.*, 1934, **41,** 93.)

For Reference

REYNOLDS (1937). The Nature of Uterine Contractility. *Physiol. Rev.*, **17,** 304.
REYNOLDS (1939). Physiology of the Uterus. (London : Hamilton.)

General References on Reproduction

ALLEN (Editor), and 26 others (1939). Sex and Internal Secretions. (Baltimore : Williams & Watkins ; London : Baillière, Tindall & Cox.)
ALLEN and DOISY (1927). Ovarian and Placental Hormones. *Physiol. Rev.*, **7,** 600.
BEACH (1947). A Review of Physiological and Psychological Studies of Sexual Behaviour in Mammals. *Physiol. Rev.*, **27,** 240.
CORNER (1943). The Hormones in Human Reproduction. (Oxford : Univ. Press.)
KENNEDY (1926). Diet and Sterility. *Physiol. Rev.*, **6,** 485.
LAQUEUR (1946). Inter-relationships between Gonadotrophic and Sex Hormones. *Harvey Lects.*, **41,** 216.
MARSHALL (1922). The Physiology of Reproduction. (London : Longmans, Green.)
NEWTON (1939). Recent Advances in Physiology. (London : Churchill.)
PARKES (1929). The Internal Secretions of the Ovary. (London : Longmans, Green.)
ROBSON (1939). Recent Advances in Sex and Reproductive Physiology. (London : Churchill.)

CHAPTER LIII

THE SECRETION AND PROPERTIES OF MILK

After birth the nutriment supplied to the child normally depends on the supply of milk by the mammary glands. Nutrition now involves further activity on the part of the infant, the alimentary canal being concerned in the digestion of the milk supplied by the mother, and the kidneys being made use of for getting rid of waste material.

Growth and Functional Differentiation of the Mammary Glands. The preparation of the mammary glands for the subsequent nourishment of the newborn child begins in the first month of pregnancy, and is marked by swelling of the glands, rapid proliferation of the duct epithelium, and production of many new secreting alveoli. As we have seen, there are probably at least three hormonal factors concerned in these processes—œstrogen and progesterone from ovaries and placenta (an optimal ratio is 240 I.U. œstrone to 1 mg. progesterone), and prolactin from the anterior pituitary. Œstrogens, progesterone and pituitary hormones are also available in the course of the normal menstrual cycle, and are no doubt responsible for the periodical enlargement of the breasts at each menstrual period. Throughout these changes, and throughout pregnancy, however, the secretory activity of the enlarged glands is held in abeyance.

The œstrogens cause an increased vascularity of the mammary tissue and enhance its permeability to progesterone : they induce growth, chiefly of the duct system, and of the epithelium of the nipple, while progesterone causes growth and differentiation, up to the threshold of lactation, in the secreting alveoli. Many steroids, and also diethylstilbœstrol, have a similar effect. Œstrogens and progesterone are effective in ovariectomised animals, and, when given in large doses, can produce growth and development of the mammary glands, and even lactation, in males. But they are ineffective in causing mammary growth in ovariectomised and hypophysectomised animals, though the effect on the nipple is obtained.

Doubt exists as to the part played by the anterior pituitary. One view is that the œstrogen and progesterone incite the pituitary to produce two new hormones, called respectively Mammogen I and Mammogen II. Others think that the effects can be attributed to a collaboration between already recognised anterior pituitary hormones, and particularly the growth hormone and prolactin, on the one hand, and œstrogen and progesterone on the other. Implantation of pituitary after hypophysectomy restores the proliferating effect, and hence collaboration between a pituitary hormone, œstrogen and progesterone seem essential at an early stage, as well as later.

The placenta can probably also play during pregnancy a part similar to that of the anterior pituitary, since if pregnant animals of some species are hypophysectomised, mammary development continues. In mice, the fœtus can be killed by crushing, but the placentæ remain till term : mammary growth then continues : but if the placentæ are also removed, the glands atrophy.*

Lactation. With the expulsion of the fœtus at birth the overgrown gland cells enter into a condition of activity, and the secretion of the gland, once formed, is continued independently of the fœtus, or, indeed, of any of the pelvic organs. For this initiation and maintenance of secretion, hormones of the anterior pituitary are mainly responsible. Certainly lactation is at once arrested after hypophysectomy,† while on the other hand, if preliminary breast growth be induced, by previous dosage with œstrogen and progesterone, injection of crude anterior pituitary extracts will cause lactation in immature females, or even in males. They have been successfully employed clinically in cases of failure of lactation.[*]

---

* Newton and Beck. *Endocrin.*, 1939, **1**, 65.

† Selye, Collip and Thompson. *Endocrinol.*, 1934, **18**, 237.

Among the anterior pituitary hormones which are responsible for the induction of lactation, the lactogenic hormone, prolactin, is important: but it has been found that, when pure, it is not by itself competent to precipitate full lactation or to maintain and reinforce a declining lactation.[†] Its effect is possibly due to an influence on carbohydrate metabolism. The presence of the adrenocorticotropic hormone, together with prolactin, is also probably essential for lactation.

Other endocrine organs also play an essential *rôle* in the determination and maintenance of lactation. The adrenals, for instance, are essential, and after their removal lactation ceases, probably because of fundamental changes in general metabolism. It has been shown that, of the various adrenal steroids, the essential one is that responsible for carbohydrate metabolism (17-hydroxy-11-dehydro-corticosterone), though there must also be present enough desoxycorticosterone to conserve normal salt and water balance.

The thyroid appears to play no indispensable part in the initiation of lactation, but contributes unquestionably to its upkeep: thyroidectomy often diminishes the milk yield, while administration of thyroxine definitely raises the milk fat content. The influence of the thyroid is probably indirect, and due to its effect on the general metabolism. Parathyroid hormone is also essential for full milk yield, which is not surprising in view of the high calcium content of milk.

An important question is why, in most cases, lactation does not commence until after parturition. One view is that lactation is held in abeyance by some such factors as the presence of the corpora lutea, or the placenta, both of which would act by formation of hormones, or by the mechanical distension of the uterus, which would presumably be reflex in nature. None of these suggestions adequately accounts for the phenomenon in all cases, or in all species. It seems more probable that by full term the mammary glands are fully developed, and only await the stimulus of prolactin, in presence of other essential hormones, for full lactation to commence. There is an increase in the prolactin content of the anterior pituitary towards term, and during the period of nursing, which is probably sufficient to initiate and maintain lactation, provided the milk is regularly drawn off.

During the latter half of pregnancy, a watery fluid can generally be expressed from the nipple. In certain mammals, this watery secretion gives place to a secretion of true milk at the end of gestation, or during the process of parturition itself. In the woman, though the formation of milk may be hastened if a child has been put to the breasts during the latter part of pregnancy, secretion normally begins on the second or third day, even if the child has been born dead and no attempt at suckling has taken place.

Removal of Milk. Normally, the only condition which is necessary for secretion to continue during six to nine months after parturition is the repeated removal of the secreted milk. The process of suckling removes the milk already formed, and this leads to the secretion of more milk. This it probably does, in part, by causing reflexly an increased formation of lactogenic hormone in the anterior pituitary.[‡] The reflex may originate from the nipples, but the release of tension on the uterine walls seems to be an accessory condition, since emptying of the uterine contents may precipitate lactation, while distending the emptied uterus with paraffin wax delays or inhibits it.

A more important factor in suckling is probably a release of oxytocin into the blood by a reflex which acts upon the posterior pituitary :[*] this causes contraction of the plain muscle of the capacious milk sinuses, and so aids in the expulsion of the milk. It is of importance, therefore, to distinguish between milk secretion and the expulsion or removal of the milk in any way.

---

[*] Riddle. *J. Amer. Med. Assoc.*, 1935, **104**, 636.

[†] Folley and Young. *Bioch. J.*, 1939, **33**, 192; *Lancet*, 1941, **p. 380**.

[‡] Selye. *Amer. J. Physiol.*, 1934, **107**, 535.

It is well known that if milking is suspended, further formation of milk soon ceases, while more frequent milking leads to production of a larger total milk yield. This appears to be because, as soon as the pressure in the mammary alveoli reaches a certain figure—some 25–40 mm. Hg.—further secretion of milk ceases : frequent emptying therefore, by keeping the pressure on the secreted milk low, encourages further secretion. An increase of blood sugar, or administration of thyroxine, cause in cows an increase in the volume and concentration of the milk.[†] If a woman does not nurse her child, the swelling of the breasts gradually passes off, the milk disappears, and the glands undergo a process of involution. Under normal conditions the secretion of milk lasts for six to nine months, and may in rare cases extend over more than a year. The amount secreted increases at first with the growth and size of the child, and often may reach 1·5 to 2 l., or occasionally 3 l. daily. The table above represents the average amount of milk secreted during the thirty-seven weeks after confinement. It will of course be greater with strong, big children, and smaller with weakly children.

Table showing Amount of Milk Secreted by a Nursing Woman

Time	Milk secreted daily
INCREASE	
1st day	20 grammes.
2nd ,,	97 ,,
3rd ,,	211 ,,
4th ,,	326 ,,
5th ,,	364 ,,
6th ,,	402 ,,
7th ,,	478 ,,
2nd week	502 ,,
3rd–4th week	572 ,,
5th–8th ,,	736 ,,
9th–12th ,,	797 ,,
13th–16th ,,	836 ,,
17th–20th ,,	867 ,,
21st–24th ,,	944 ,,
25th–28th ,,	963 ,,
DECREASE	
29th–32nd week	916 grammes.
33rd–36th ,,	909 ,,
37th week	885 ,,

Colostrum. Before the secretion of true milk begins, the fluid which is obtained from the breasts is known as colostrum. It may be expressed from the breasts immediately after delivery and is ingested by the child during the first two days. The colostrum is formed only in slight quantities. It is an opalescent fluid, often somewhat yellowish, containing fat globules which, if the fluid be allowed to stand, form a yellowish layer on the top. Under the microscope, in addition to the fat globules, may be seen the so-called colostrum corpuscles, which are multinucleated cells loaded with particles of fat. They are probably leucocytes or phagocytes which have wandered into the alveoli and have taken up fat globules. Some of the corpuscles may be desquamated secretory cells. Colostrum is distinguished from true milk by containing little or no caseinogen. It contains about 8 per cent. of proteins, namely, lactalbumin and lactoglobulin, which coagulate on boiling. Lactose and fat are present in smaller, and salts in greater amount than in ordinary milk. It is popularly supposed to have a laxative effect on the child.

A fluid resembling colostrum is obtained at later stages if withdrawal of milk is suspended for a few days. The explanation of colostrum is that it is formed by retention of the milk as soon as the intra-alveolar pressure reaches a level at which secretion ceases, the contents tend to come into equilibrium with the blood, and the resultant fluid is colostrum.

---

[*] Ely and Petersen. *J. Dairy Sci.*, 1941, **24**, 211.

[†] Folley and White. *Proc. Roy. Soc.*, 1936, **120B**, 346.

THE SECRETION OF MILK. When fully formed, each mammary gland consists of fifteen to twenty lobes embedded in connective tissue. Each obe is made up of a mass of secreting alveoli, which lead by narrow ducts into one large lactiferous duct. These lactiferous ducts, one from each lobe, open on the nipple, undergoing in the nipple itself an oval enlargement. Before pregnancy, the alveoli as well as the ducts are lined with a cubical epithelium. Early in pregnancy a marked difference develops between the epithelium of the alveoli and that of the ducts.

The secreting cells increase in size and especially in height, and numerous small fat cells, mitochondria and spherical "pseudo-yolk" granules are present.* At later stages the fat globules increase in size and collect towards the lumen, while the pseudo-yolk granules disappear. When lactation sets in the cells contain elongated mitochondria and one or two large fat globules next to the lumen; ultimately the fat globules are discharged, together with a little cytoplasm, by a budding-off process, into the secretion in the lumen (Fig. 689). During the stage of enlargement and also during lactation there are frequent mitotic divisions in the alveolar cells. A few cells degenerate and are shed, as also are a certain number of leucocytes.

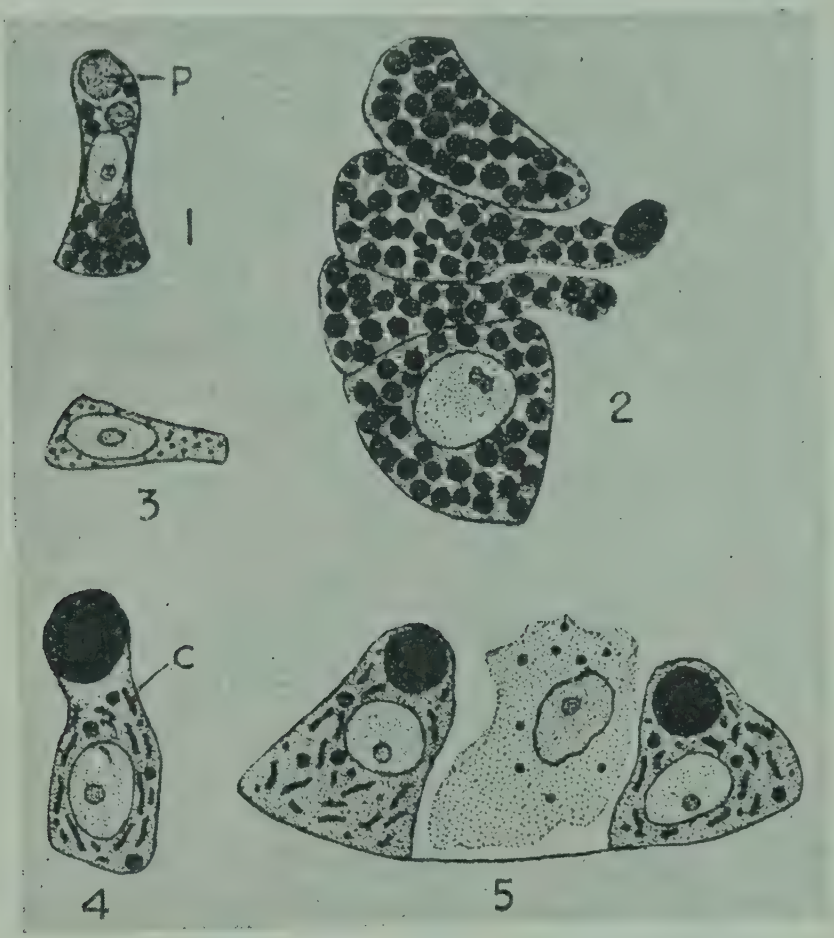


FIG. 689. Cells of Mammary Gland of Rat. (1) and (2) On eighteenth day of pregnancy: Fig. 2 shows elimination of a large fat drop. (3) A cell after fat has been secreted, a few hours before parturition. (4) Cell in full lactation, seven days after parturition. (5) Two normal cells and one that is degenerating and passing into the lumen; twelfth day of lactation. p = pseudo-yolk spheres, c = chondriosomes. (JEFFERS. *Amer. J. Anat.*, 1935, **56**, 257.)

The secreting mammary gland is very vascular, about 300 to 400 volumes of blood being needed for each volume of milk produced.

We know very little about the biochemistry of milk secretion. It seems impossible at present to explain the very close adaptation between the activity of the secretory cells and the needs of the infant or young animal. Two at least of the constituents of milk, caseinogen and lactose, are peculiar to this secretion. It would seem that the caseinogen is produced by the gland cells by synthesis from the amino acids and probably also from globulin of the blood stream, and that the lactose is derived, also by synthesis, from the blood sugar, and perhaps also from lactate. Slices of surviving lactating mammary gland can form lactose from glucose, but not from fructose, mannose or galactose.† The fat similarly originates in the blood fat,‡ and much of it ultimately in the fat of the food, or of the body depots. The phosphorus compounds of milk are derived from the inorganic phosphate of the blood.

---

* KATHARINE R. JEFFERS. *Amer. J. Anat.*, 1935, **56**, 257.

† GRANT. *Bioch. J.*, 1935, **29**, 1905; 1936, **30**, 2027.

‡ GRAHAM, JONES and KAY. *Proc. Roy. Soc.*, 1936, **120B.**, 330.

The food of the mother may be varied within very wide limits without altering the composition or amount of the milk secreted, provided that the food is sufficient in amount. The only constituent of the milk for which we have direct evidence of alteration by changes in the food supply of the mother is the fat. It is well known that the composition of butter may be affected according to the food supplied to the cow. A large supply of oilcake, for instance, may result in the production of a butter which is deficient in the saturated fatty acids and is therefore oil at ordinary temperatures. Abnormal fats and fatty acids such as iodised fats or erucic acid, when administered to an animal in lactation, may appear among the fats of the milk.

Not only can the secretion and composition of the milk be affected reflexly through the nervous system, as *e.g.* under the influence of emotions, but the influence may be reciprocal. This is especially marked in the case of the pelvic organs. The act of suckling excites rhythmic contractions of the uterus (*v.* Fig. 688). Putting the child to the breast shortly after birth is, therefore, an important means of causing contraction of the uterus, and stopping any tendency to hæmorrhage from the venous sinuses opened by the separation of the placenta and fœtal membranes. The nursing of the child is thus an important means of procuring a proper involution of the uterus after labour.

PROPERTIES OF MILK

Fully formed milk presents certain features which are common to all mammals. We may therefore deal with the composition of cow's milk and point out later in what respects human milk differs therefrom. Milk is an opaque, white fluid, with characteristic odour and sweetish taste. Its specific gravity varies between 1028 and 1034. The pH = 6·6—6·8, and it therefore contains no bicarbonate ; as it contains buffers, *e.g.* phosphates and caseinogen, it may be titrated, with suitable indicators, against either acids or alkalies. One hundred cubic centimetres of average fresh milk, when titrated to methyl orange, require 41 c.c. $n/10$ acid for neutralisation. When titrated to phenolphthalein the same amount requires 18 c.c. $n/10$ alkali for neutralisation. When exposed to the air, milk rapidly undergoes changes in consequence of infection by micro-organisms. The most common of these changes is the formation of lactic acid by the *bacillus lacticus.* Exceptionally, the milk may undergo a species of alcoholic fermentation, as in the formation of *kephir,* which is made by the fermentation of mare's milk.

Composition of Milk. FATS. The opaque appearance of milk is due chiefly to the presence of multitudes of fine particles of fat. On allowing the milk to stand, or by centrifugation, the particles rise to the surface, forming cream, and by mechanical agitation, especially if the milk is slightly sour, they may be caused to run together with the formation of butter. The reason why the fat globules do not run together naturally is probably because, in any protein solution, such as that in which the globules are suspended, the protein tends to aggregate, with the formation of a pellicle, at the surface, so that an emulsion once produced in such a fluid will tend to be more or less permanent. The fats of milk consist for the greater part of the neutral glycerides, tripalmitin, tristearin, and triolein. In smaller quantities it contains the triglycerides of butyric acid, and all other saturated fatty acids, of even carbon numbers up to lignoceric (C_{24}), and various unsaturated acids also. It also contains some lecithin and cholesterol.

The milk plasma, the fluid in which the fat globules are suspended, contains various proteins, lactose, and inorganic salts, with a small amount of lecithin and nitrogenous extractives.

PROTEINS. The chief protein of cow's milk is *caseinogen*, a phosphoprotein soluble in dilute alkalies. It may be prepared from separated milk by the addition of weak acids. The precipitate may be washed by decantation, and purified by solution in dilute ammonia and precipitation by acetic acid two or three times. The precipitate is extracted with alcohol and ether, and forms a snow-white powder which is practically insoluble in water and dilute salt solutions. It is easily dissolved on the addition of a little alkali, when it yields solutions which are acid to litmus. When rubbed up with chalk it dissolves, displacing carbon dioxide and forming a calcium caseinogenate. A solution of caseinogen in NaOH or KOH is transparent and passes easily through a clay cell. The calcium caseinogenate forms opalescent solutions. Apparently the compound is hydrolysed by water with the formation of caseinogen acid, which is in a state of partial solution as swollen-up aggregates. It is impossible, therefore, to filter calcium caseinogenate through a clay cell. It is mainly in this form that caseinogen is contained in milk, hence the opalescent appearance of the milk plasma. On treating caseinogen with chymosin, it is converted into a modification known as *paracasein*, which in the presence of calcium salts is thrown out as insoluble *casein*. To this process is due the clotting of whole milk by rennet, which is made use of in the preparation of cheese, the curd consisting of a spongework of casein enclosing fat globules in its meshes. On allowing the clot to stand it shrinks, pressing out the whey.

From the whey may be obtained two other proteins, known as *lactalbumin* and *lactoglobulin*. These resemble the albumin and globulin of blood serum. They are coagulated on heating, and form a pellicle on the surface, as milk does. In human milk there is more lactalbumin (about 0·7 per cent.) than caseinogen (about 0·3 per cent.). According to some authors a third protein is present in the whey, to which the name whey protein has been given, and which is supposed to be split off from the caseinogen under the action of chymosin.

Milk can be boiled without undergoing general coagulation. If it be allowed to stand and become slightly sour by the formation of lactic acid, at a certain degree of acidity boiling the milk causes its complete coagulation. Later on the acid produced is sufficient in itself to precipitate the caseinogen. Both these processes—namely, coagulation of half-sour milk by heating, and spontaneous clotting of milk by the production of acid—are made use of in different countries for the manufacture of cheese.

Carbohydrate. The lactose is most easily obtained from whey, which is boiled to precipitate the remaining proteins. On filtering and evaporating slowly, the milk sugar crystallises out. Lactose is a disaccharide which on hydrolysis by acids, or by the lactase of the intestinal juice, yields glucose and galactose. It may be found in the urine of nursing women when the breasts are not kept empty, owing to reabsorption into the blood of the lactose formed in the mammary glands.

Salts. The salts of milk include calcium, sodium and potassium as phosphates, citrates, proteinates and chlorides. It is said to be deficient in iron.

Significance of Constituents. The composition of milk depends on many factors, such as diet, health, frequency of milking, exercise, emotional states, etc. Human milk shows a change in composition in the course of lactation, as shown in the Table below, the protein declining, the lactose and fat increasing slightly.

Average Composition of Human Milk (After Bell[*])

Time.	Protein.	Lactose.	Fat.
5 days	2·00	6·42	3·2
9 ,, ...	1·73	6·73	3·7
3–4 weeks	1·37	7·11	3·6
5–6 ,, ...	1·30	7·11	4·0
7–8 ,, ...	1·21	7·11	4·0

Milk contains vitamins A, B_1, B_2, C and D in adequate amount.

[*] Bell. *J. Biol. Chem.*, 1928, **80**, 239.

Milk is excellently adapted to the needs of the growing animal. In the first place, we find a proportionality between the total solids of the milk and the rate at which the young animal grows. The milk taken by the animal serves only in part for the production of energy, a great proportion of it being required for the building up of new tissue. In no respect is this correspondence seen better than in the comparative analyses of the ash of milk and of the young animal of the same species. Thus, the calcium content of human milk is 0·04 per cent., and that of rabbit's milk about 0·9 per cent. This close correspondence is necessary only where the growth is very rapid, so that the greater part of the constituents of the milk have to be utilised in the building up of the animal tissues. The slower the growth of the animal the greater the divergence between the composition of the milk and that of the newborn animal.

The relation between rate of growth and protein content of the milk is well illustrated by a comparison of the composition of the milk in different animals. In the following Table (Proscher) it will be seen that the more rapidly an animal grows the greater is the protein content of the milk supplied by its mother :

	Time in which the body weight of the newborn animal was doubled. Days.	Protein. Per cent.
Man	180	1·6
Horse	60	2·0
Cow	47	3·5
Goat	19	4·3
Pig	18	5·9
Sheep	10	6·5
Dog	8	7·1
Cat	7	9·5

We should expect that the milk, which is the sole food of the growing infant, would contain a relatively greater proportion of protein than is necessary for the adult. In an experiment by E. Feer, quoted by Bunge, a child weighing 8,226 g. at the thirtieth week took 951 g. of milk daily. Human milk contains :

Protein	1·6 per cent.
Fat	3·4 ,,
Sugar	6·1 ,,
Ash	0·2 ,,

The child was therefore receiving daily :

Protein	15·4 g.
Fat	32·3 ,,
Sugar	58·0 ,,
Ash	1·9 ,,

According to the same proportions a man of 70 kilos. fed entirely on milk would take in

Protein	129 g.
Fat	275 ,,
Sugar	494 ,,
Ash	16 ,,

It is interesting to note that the protein of this diet is somewhat above that in the diets ordinarily accepted as standard, but there is a large excess in the fat and in the total Calorie value, as would be expected from the more rapid metabolism and the relatively larger body surface of the young child.

The fitness of caseinogen for building up the tissues of the body is evident when we compare the products of its hydrolysis with those of the proteins in other foodstuffs. It will be seen from the Table on p. 44 that practically every amino-acid and allied substance employed in the building up of the various proteins is represented in caseinogen. The only exceptions are glycine, which can easily be synthesised in the body, and cysteïne, which is present, however, in lactalbumin.

We thus see that, under normal conditions, the young animal is supplied, through its natural food, with all the foodstuffs in the proportions which it requires for its normal nourishment and growth. It is not easy, however, satisfactorily to replace the natural milk of an animal by that of another species. In civilised communities it is becoming more and more the custom to endeavour to feed infants with cow's milk, more or less modified, in the vain endeavour to reproduce the properties of human milk. Among all classes, this involves the administering of a milk differing in its properties and in the relative proportions and qualities of its proteins, fats, carbohydrates, and salts, from human milk. So-called "humanised" milk is only a rough imitation of the natural mother's milk. Among the poorer classes this artificial feeding means the replacement of a natural sterile food, throwing very little work on the digestive organs of the child, by a foreign milk, very difficult to digest and often teeming with micro-organisms. In some cases it is necessary to adopt artificial feeding because there is an insufficient secretion of milk. It is therefore important to know what are the main differences in composition between human and cow's milk. In human milk, the caseinogen is, not only absolutely, but also relatively, less than in cow's milk, while the latter is relatively poorer in lactose. Human milk is poorer in salts, especially in lime, containing only one-sixth of the amount present in cow's milk. The main difference may be summarised as follows :—

	Water.	Proteins.		Fat.	Lactose.	Salts.
		Caseinogen.	Albumin			
Human milk . .	88·5	0·3	0·7	3·7	6·7	0·2
Cow's milk . .	87·1	2·3	0·7	4·0	4·8	0·7

The simplest modification of cow's milk to adapt it for the human infant is to dilute it, and add cream and lactose.

For Reference

Davies (1936). The Chemistry of Milk. (London : Chapman & Hall.)
Espe (1941). Secretion of Milk. (Ames, Iowa : Collegiate Press Inc.)
Meigs (1922). Milk Secretion as Related to Diet. *Physiol. Rev.*, **2**, 204.
Petersen (1944). Lactation. *Physiol. Rev.*, **24**, 340.

CHAPTER LIV

HEREDITY [*]

FROM the fertilised human ovum, there is formed, as a result of some fifty successive mitotic divisions, an individual consisting of about 2^{48} cells,[†] and partaking of many of the minutest characteristics of both its parents. An egg cell has much the same appearance, whether it belong to an echinoderm, a fish, or a woman. In the process of development, by a simple repetition of a series of cell divisions, this undifferentiated protoplasm is formed into the complex organs with the potentialities and habits which distinguish the type from which it has been derived. There must, therefore, be some difference between different eggs. We cannot wonder that the intimate nature of this process has been the subject of speculation from the very birth of science. Running through these speculations were two main ideas, which have been labelled the theories of "evolution" and of "epigenesis." By the "evolutionists" the egg was believed to contain an embryo fully formed in miniature, as the bud contains the flower or the chrysalis the butterfly. Development, therefore, was only the unfolding of something already existing. If, however, this theory be pushed to its utmost, and if the egg contain a complete embryo, this must itself contain eggs for the next generation, and so on *ad infinitum*, a conclusion which of course is absurd. According to the theory of epigenesis, the structure of the egg is wholly different from that of the adult, its development into the new individual consisting in the continual formation one after the other, of new parts previously non-existent as such. There is no doubt that this view is fundamentally correct.

The difficulty with which we have to contend is the understanding of the orderly sequence and correlation of the cell divisions and differentiations which result in an adult individual of the same type as the parents. The fact that, under approximately identical conditions, one mammalian ovum will give rise to a mouse and the other to a man, indicates that there must be some preformed differentiation in the nuclear structures, *i.e.* chromosomes, of the ovum itself. We have already seen that in some types the germ cells are separated off from the rest of the embryo at a fairly early stage. If, in the two-celled stage of the frog's egg, one cell be destroyed by means of a hot wire, the other cell develops to form half an embryo, thus suggesting that each cell of the two-celled embryo could give rise only to the corresponding half of the body. This limitation of development, however, occurs only if the intact cell be left in connection with the cell that has been injured. If, in echinoderm larvæ, the cells be entirely separated, even as late as the eight-celled stage of division, each cell will give rise to a whole embryo, differing from a normal embryo only in respect of size. In man too, "identical twins" result similarly from separation of the two cells formed by the first division of the fertilised ovum. This suggests that the number of divisions that each cell can undergo is predetermined in the egg itself, but shows also that the cells into which the egg divides are, at first at any rate, equipotential. We must assume, therefore, that the reason why one cell under ordinary circumstances forms only one half of the embryo is that its development is regulated and determined by the presence of the other cell in connection with it and forming the other half of the embryo. That is to say, the development of the egg involves the adaptation to environment of a protoplasm of certain properties and powers of eaction. The final product of development depends (1) on the nature of the protoplasm (including the nucleus) of which the egg is composed, and (2) on the environmental conditions to which the egg is subjected during the rapid growth and multiplication attending its development. We could therefore speak of a morphology of inheritance, but the morphology would be ultra-microscopic and have relation to the molecular structure of the protoplasm of which the egg was composed.

---

[*] The subject of genetics is now highly specialised, and it must be understood that all that can be attempted here will necessarily be of the nature of a very superficial synopsis of salient points.

[†] HALDANE. *Proc. R. S.*, 1948, **135B**, 168.

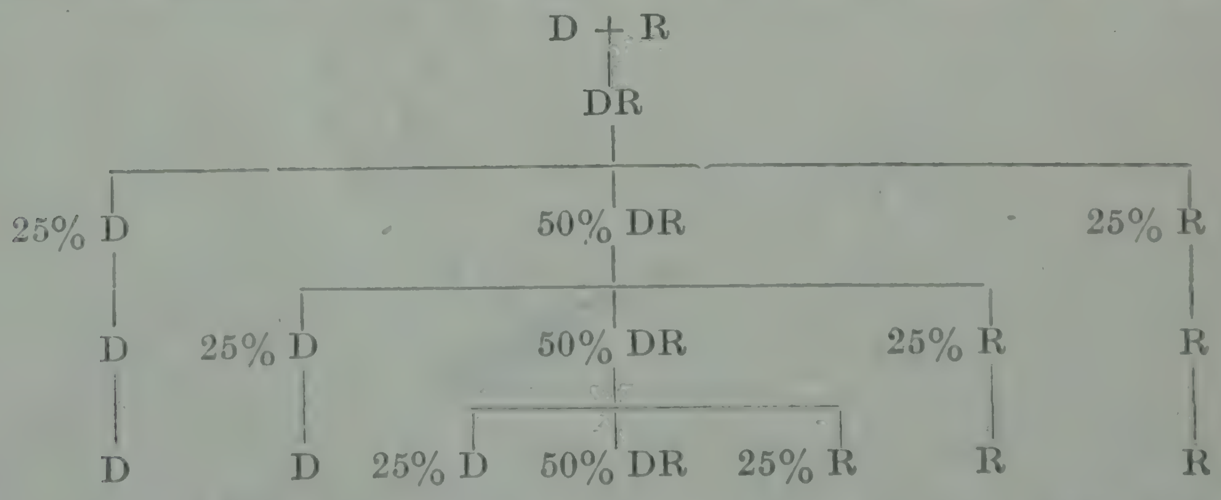
In the transmission of the potentialities of development from parent to fertilised egg, we must regard the nucleus as the essential structure. In ordinary development the spermatozoon furnishes only half a nucleus and a centrosome, the ovum supplying the other half nucleus and the whole of the cytoplasm. In division of the egg, the only part of the cell which can divide so that each daughter cell shall include an equal part of both parental germs is the nucleus. The constant number of the chromosomes in each species, and their accurate division on mitosis, suggest that the hereditary transmission of the potentialities of the cell is bound up with the chromosomes. It is now believed that every inheritable character is located in a chromosome, or part of a chromosome. If this be the case, one might regard the differentiation into various tissues, which occurs in the process of development, as occasioned by the modification or actual loss of the constituent parts of one or more of the chromosomes present in the pluripotential fertilised egg-cell. There is no doubt that many tissues do become thus differentiated at a fairly early period in development, having undergone in the process an absolute modification of their potentialities, which must be at any rate shared by the chromosomes of their constituent cells. The extent to which this limitation of powers of development occurs varies widely in different animals. In the higher animals, such as man, epithelium will reproduce epithelium and liver cells will reproduce liver cells, while nerve cells once definitely characterised are absolutely incapable of multiplication. On the other hand, in crustacea, a whole limb may be torn off and be regenerated from the tissues of the stump. Destruction of the optic lens in the salamander is followed by its regeneration from the anterior part of the optic cup, a tissue which had no part in its primary formation. Worms will form a new head after decapitation. In these animals, therefore, the cells in many parts of the body possess the power of directed growth, if need arise in case of injury, and are able to form tissues of many different kinds.

We have mentioned the small size of the larvæ formed from isolated cells of the segmenting egg as a proof that the number of cell divisions of the somatic part of the developing animal is predetermined and limited. This conclusion must not be taken too absolutely. Many of the tissues, even of the highest animals, possess the power of almost unlimited regeneration by cell multiplication as a response to injury. Under normal conditions, the growth of such tissues is limited, not by absence of power to divide, but as a result of a mutual interaction between them and the surrounding cells. We might almost speak of a struggle for existence between the various tissues of the body, which, in the healthy organism, results in a balance of multiplicative powers. If this balance is upset by any means, such as stimulation of certain cells by the presence of intracellular parasites, or their destruction by irritants or other abnormal conditions (*e.g.* exposure to X-rays), one tissue may enter into active growth at the expense of the surrounding tissues, and the result is a morbid growth such as cancer. It is possible that in the latter case a new factor comes into play. All tissues of the body, as we have seen, begin to die from the time that they are born. They have a certain span of life, a certain limitation to their generations, which results in the phenomenon of senescence, such as occurs in a culture of protozoa. In protozoa this phenomenon is the signal for rejuvenation by conjugation. It is possible that, in cancer, something of the same nature occurs. It is, at any rate, significant that in a rapidly growing cancer, many of the dividing cells present the phenomenon of meiosis, a phenomenon which is otherwise found only in the ripening sexual cells. Given adequate conditions of nutrition, there seems to be no limit to the growth of cancer cells. In mice, a cancer may be transferred from one individual to another by inoculation, and this process may apparently go on indefinitely.

In sexual reproduction, the new individual partakes of characteristics of both its parents. It therefore resembles neither of its parents in all details. The conjugation of the two parent cells, from which it is derived, has been preceded by a throwing out of half the chromosomes from each parent cell. It is therefore natural to ascribe the variations, which occur among the members of one family, to a qualitative difference in the chromosomes which have been retained in the formation of their respective egg cells. Can we regard the chromosomes as representing separate qualities of the individual, or must we assume that all qualities are presented to a greater or less extent in every chromosome? In the case of many qualities, especially those which distinguish the species as apart from the individual variation or family characteristic, we must probably accept the latter idea as correct. In this case, the child can be regarded as representing an arithmetical mean of both its parents. In certain respects, however, a quality seems to be transmitted from parent to offspring, either completely or not at all. This is specially applicable to those characters which have been rapidly produced by artificial selection, characters which, if artificial selection be abandoned, rapidly disappear, with reversion to the type from which the special strain was initially produced. These characters are alternative or *allelomorphic*— *e.g.*, tallness or shortness, blue eyes or brown eyes.

LAWS OF INHERITANCE

Mendel's First Law. The way in which these characteristics are transmitted was first studied by Mendel and has been formulated as "Mendel's laws". *The first law is* that of the *segregation of characters*, and is illustrated by his first experiments, which were carried out on peas. On crossing a tall plant with a dwarf plant, seeds were obtained all the plants from which were tall. On recrossing the plants of this generation among one another, a third generation (F2) was obtained, in which 25 per cent. of the plants were dwarfs and 75 per cent. were tall. Crossing the dwarf plants among themselves led to the production of dwarf plants through successive generations. Of the 75 per cent. tall plants one-third, and all their descendants, continued to produce tall plants when self-fertilised, whereas of the remaining two-thirds of the tall plants 25 per cent. produced dwarfs and 75 per cent. produced tall plants. On continuing the process of breeding, the dwarf plants, when self-fertilised, always produced dwarfs, whereas of the tall plants 25 per cent. produced tall plants which bred true, while the remaining 50 per cent. produced the same percentage of tall and dwarf as in the preceding generations. Mendel explained these results by the assumption that a character could be either *dominant* or *recessive*. If both characters were present together in one plant it would partake of the dominant type; the fact that this plant possessed the recessive character would be shown only by the result of breeding. In the case of the peas, the tall character was dominant over the dwarf. Thus when the tall and dwarf pea were crossed, the first generation of plants would exhibit the dominant character and be tall. In the second generation, however, 25 per cent. of the individuals would be pure dominants (D + D), 25 per cent. would be pure recessive (R + R), while 50 per cent. would be mixed (D + R). The pure dominants bred together would always give rise to nothing but pure dominants, the recessive to recessive, while the mixed type would always, as before, give rise to 25 per cent. pure dominants, 50 per cent. mixed, and 25 per cent. pure recessives. These results may perhaps be made clearer by the Table below :—



It has been suggested that a very large number, if not all, of the characters of an individual might be brought under this law. This might be done by indefinitely subdividing the characters, but the question would then become beyond the limits of analysis or experimental investigation. There is no doubt that many qualities are subject to Mendel's law, and their study has been of considerable assistance in guiding the efforts of our breeders and horticulturists in the formation of new varieties desirable for their value to man.

In respect of many qualities, the Mendelian idea of dominance seems to fail. Thus in man, the progeny of a cross between a white and a black race are more or less intermediate between the two, and vary according to the amount of black and white blood introduced in succeeding generations. Definite black and white individuals are not produced, but merely individuals of various degrees of brownness. This is known as *blending*.

Genes. The explanation now given of Mendel's first law is a simple one. The hereditary properties are comprised within the chromosomes, the chromatin of which consists of long molecules of nucleoprotein, arranged in a definite molecular pattern, which is reproduced with almost unerring accuracy when new cells are formed. This is because the newly formed chromatin is laid down on that from the parent cells, which, so to speak, provides a ready-made molecular mould, or pattern, to which the new material must conform.

Special groupings or molecules, consisting of varieties of desoxy-ribose nucleic acid, retain their position in the pattern and their configuration, and these definitely located groups are called *genes*. These, of course, are paired in the somatic cells, but single in the gametes. Each gene has a definite place in its chromosome, and in some simple instances chromosome maps have been drawn up, indicating the locations of the various genes. The human chromatin probably contains thousands of genes, each endowed with highly specific potentialities. When the two genes of an allelomorphic pair are identical in the somatic cells, then the single factors in each gamete are identical, and the condition is called *homozygous :* if they are different, it is called *heterozygous*, in which case half the gametes will have one of the allelomorphic factors and the other half the other one. Suppose we take two characters T and t, representing tall and short respectively, and mate together a gamete from a pure tall plant and one from a short plant. The somatic cells of these parents are TT and tt and the gametes T and t respectively. The fertilised cell will be Tt and the new plant from this will produce gametes, half of which, owing to their formation by reduction division, will be T and half t. If these are mated together, then by the law of chance we should get the combinations TT, Tt and tt in the ratios 1 : 2 : 1 (3 dominants to 1 recessive), and, of these, the first and last being homozygous would breed true, while the second being heterozygous would, on subsequent propagation, segregate in a similar way into pure and mixed types.

Mendel's second law relates to the *independent association of genes*, and states that the various possible allelomorphic pairs are independent of each other in inheritance. For instance, if we consider two factors Aa and Bb, in which A and B are dominant ; and cross, say, an individual who is BBaa with one who is bbAA, then the gametes of the first will be Ba and of the second bA, and the somatic cells resulting will be Bb Aa. These will produce gametes of types BA, Ba, bA and ba, in equal numbers, and when they are crossed, in accordance with the laws of chance, the offspring will be of the types **BBAA, BBAa, BbAA, BbAa,** BBaa, Bbaa, bbAA, bbAa, and bbaa, in the ratios 1 : 2 : 2 : 4 : 1 : 2 : 1 : 2 : 1 or, in every sixteen :

9 with both dominants A and B.
3 with one dominant (A) and one recessive (b).
3 with one dominant (B) and one recessive (a).
1 with both recessive characters (a and b).

The 9 : 3 : 3 : 1 is obtained obviously by the combination of two 3 : 1 ratios of dominant to recessive characters, *e.g.* in peas, A = tall, a = short, B = purple flowers, b = white flowers. If we cross a tall purple with a short white, the result will be all tall purples, but if these are self-pollinated we get :—

$$3:1\begin{cases}12\text{ Tall} = 9\text{ purple} + 3\text{ white,}\\ \;\;4\text{ Short} = \underbrace{3\text{ purple} + 1\text{ white,}}_{3\,:\,1,}\end{cases}$$

or, rearranged :—

$$3:1\begin{cases}12\text{ Purple} = 9\text{ tall} + 3\text{ short,}\\ \;\;4\text{ White} = \underbrace{3\text{ tall} + 1\text{ short,}}_{3\,:\,1}\end{cases}$$

By similar independence, where three allelomorphic characters are considered, we get various combinations of dominant and recessive, in accordance with the laws of chance, with ratios 27 : 9 : 9 : 9 : 3 : 3 : 3 : 1. The possible variations in crossings where an indefinite number of characters are considered is thus limitless.

The Determination of Sex

In many animals it is known, and in all it is probable, that one of the somatic chromosome pairs is different in the two sexes, one sex having a certain pair of chromosomes, the "sex chromosomes," which are represented by only one, or by one complete and one vestigial chromosome in the other sex, and there is good reason to suppose that the initial determination of sex is connected with this difference.*

When the reduction of the chromosomes takes place during gametogenesis and the individuals of each pair of chromosomes pass into different germ cells, one complete sex chromosome passes into each germ cell, in the sex that has two similar chromosomes, called XX (female in mammals); in the other sex where there is only one chromosome similar to these and the other one is either rudimentary (XY type) or lacking (XO), then one of the gametes will receive the X-chromosome, and the other one either no corresponding chromosome, or else the modified Y-chromosome. Thus the sex which is of XY or XO type (male in mammals) produces gametes of two kinds, while the gametes of the sex with the complete pair of sex chromosomes are of one type only. Thus when fertilisation takes place the original dimorphism of the zygotes is maintained. The chromosome mechanism at fertilisation is shown diagrammatically in Fig. 690.

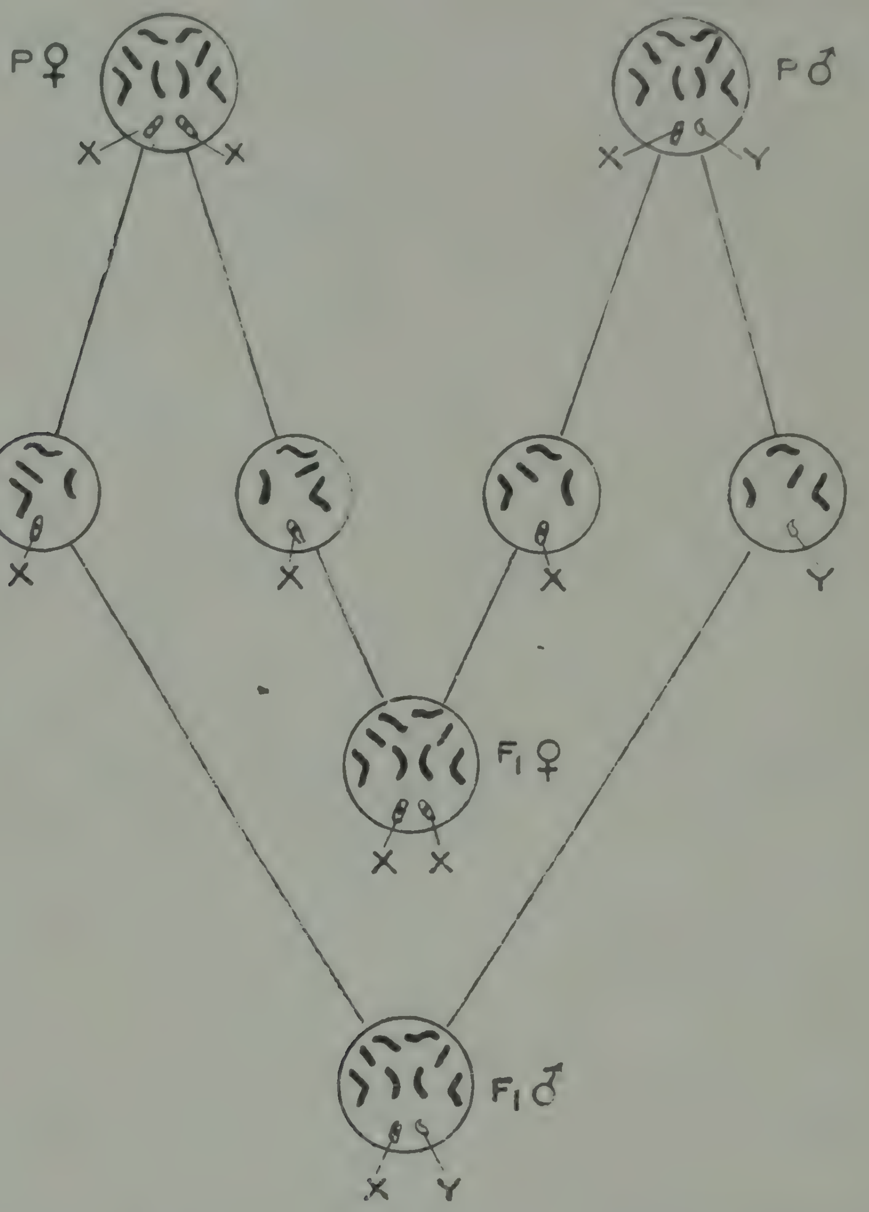


Fig. 690. Diagram to illustrate Principle of Sex-determination in an animal with 10 Somatic Chromosomes (5 pairs). P♀, Female somatic cell. P♂, Male somatic cell. G, The two ripe ova (both X), and two ripe spermatozoa (one X and one Y). F_1♀, Fertilised ovum which forms a female. F_1♂, Fertilised ovum which forms a male.

The sex which produces dimorphic gametes is therefore "heterozygous" for sex, while the other is "homozygous." The study of the correlation between sex and chromosome constitution has been greatly facilitated by the study of sex-linked characters, such as hæmophilia, their mode of inheritance lending strong support to the supposition that sex is initially dependent on chromosome constitution.

The sex carrying the extra chromosome varies in different groups of animals. In Lepidoptera and birds, for instance, the female is heterozygous and produces two kinds of ova, while the spermatozoa are homozygous, and in this case, therefore, the sex of the future embryo depends on the type of ovum fertilised. In Diptera and mammals (including man), on the other hand, the ova are all of the same type, but there are two kinds of spermatozoa. In the now well-known case of the fruit fly, Drosophila, there are four pairs of chromosomes, comprising three pairs of ordinary chromosomes in both sexes, together with an XX pair in the female, and an X and Y in the male (Fig. 691). The spermatozoa are, therefore, of two types, having three ordinary chromosomes and an X, or three ordinary chromosomes and a Y. Ova fertilised by the former produce females, and those fertilised by the latter produce males. In the human subject the male again is

---

* McClung. *Biol. Bull.*, 1902, 3, 43; Winiwarter. *Arch. de Biol.*, 1938, **49**, 119.

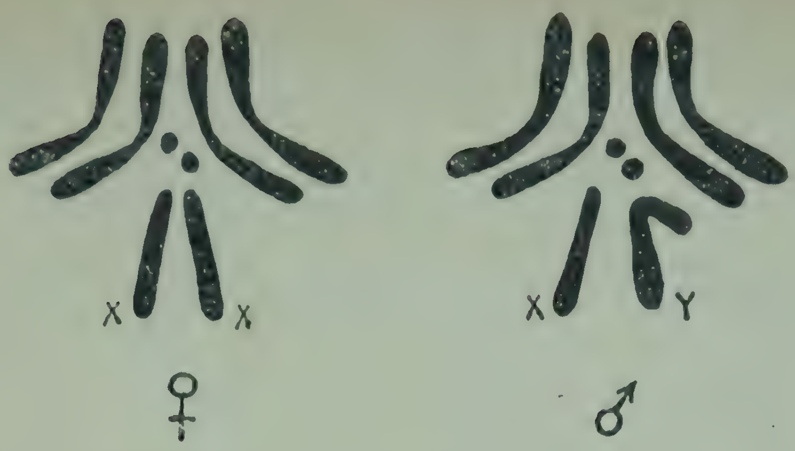


FIG. 691. Female and Male Groups of Chromosomes (Diploid) of *Drosophila melanogaster*. (After MORGAN, in Cowdray's *General Cytology*.)

heterozygous, the somatic number of chromosomes (according to Painter, Evans and Swezy, and others) being 46 + XX in the female and 46 + XY in the male. The Y chromosome is the smallest. The spermatozoa, therefore, contain 23 + X or 23 + Y, so that fertilisation of ova (23 + X) by the former results in females, and by the latter, in males. There is little doubt that, normally, the modelling of the embryo to the male or female pattern is primarily dependent upon its chromosome constitution, but there is equally little doubt that, in rare cases, other factors may lead to the chromosome constitution being over-ridden, causing the embryo to be partially or entirely diverted to the sex to which it was not predestined, and resulting in the production of intersexes or sex reversals. The underlying cause for such change is most often due to exposure of the embryo to hormones characteristic of the opposite sex. Thus in cattle, the free-martin is a masculinized female which while *in utero* was exposed to hormones from a male twin.

In man, *pseudohermaphrodites* occur, both male, *i.e.* in which the gonad is a testis, and female, *i.e.* in which the gonad is an ovary. The external genitalia and secondary sex characters approximate more or less accurately to those of the opposite sex. The male type is probably due to the influence of female hormones derived from the maternal circulation, the female type to the effect of excessive adrenal cortical hormones.

True hermaphrodites, possessing both types of gonads, are extremely rare.

Linkage

Mendel's second law is not of universal application, owing to the existence of what is termed linkage. Experiment shows that certain factors are inseparable or only partly separable from certain others. Some factors are sex-linked, a good example being in *Drosophila*. When a white-eyed male is mated to a wild red-eyed female, all the first

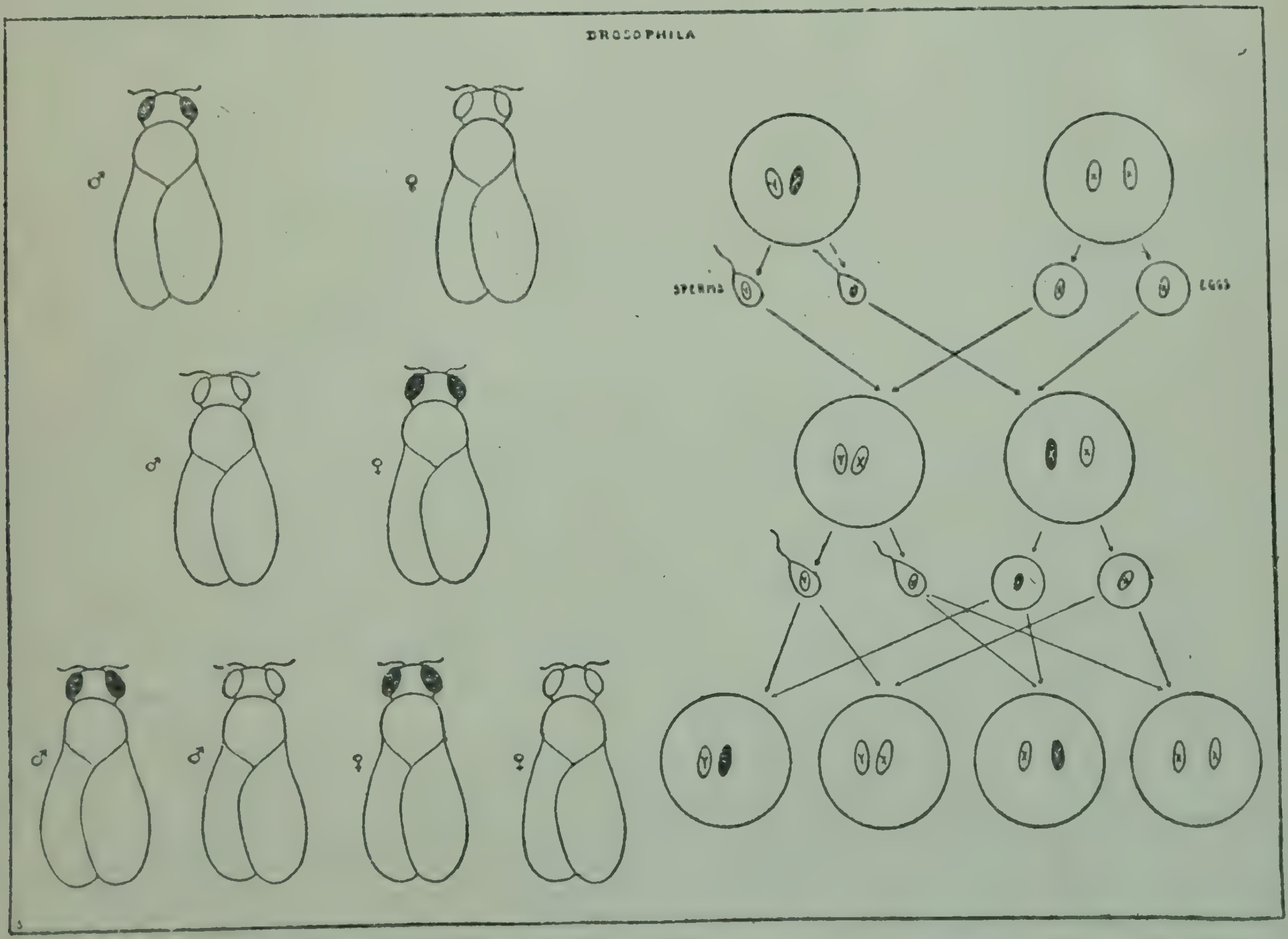


FIG. 692. Sex-linkage in *Drosophila*. Three successive generations at left: red eyes shown in black. The history of the sex-chromosomes through three generations shown at right. X-chromosome of original male in black. (SHARP, after MORGAN.)

generations (F_1) are red-eyed. When these are inbred, the F_2 consist of three red eyes to one white, all the latter being males. Consideration of these results indicates that the eye-colour factor is transmitted by the X-chromosome but not by the Y. Representing (WX) as the chromosome with the linked gene, red + sex, and (wX) as = white + sex, we have in this case :—

P_1. (wX)Y × (WX) (WX)
(white-eyed male) (red-eyed female)
F_1 = (WX)Y : (WX) (wX) (all red-eyed)
(male) (female)
F_2 = (WX) (WX) : (WX) (wX). 50 per cent. females ; all red-eyed,
and (WX)Y, 25 per cent. red-eyed males,
and (wX)Y, 25 per cent. white-eyed males.

When, as a fresh experiment, a red-eyed male is mated with a white-eyed female, then we get all the F_1 males white-eyed and all the F_1 females red-eyed.

This case is illustrated pictorially in Fig. 692.

Examples of sex-linked inheritance in man are colour-blindness and hæmophilia; these are borne by the X-chromosome and are recessive characters, hence they are carried by the female to a portion of the male descendants, but the females do not themselves suffer from the condition, unless, very exceptionally, both their parents were affected, or the father affected and the mother a carrier.

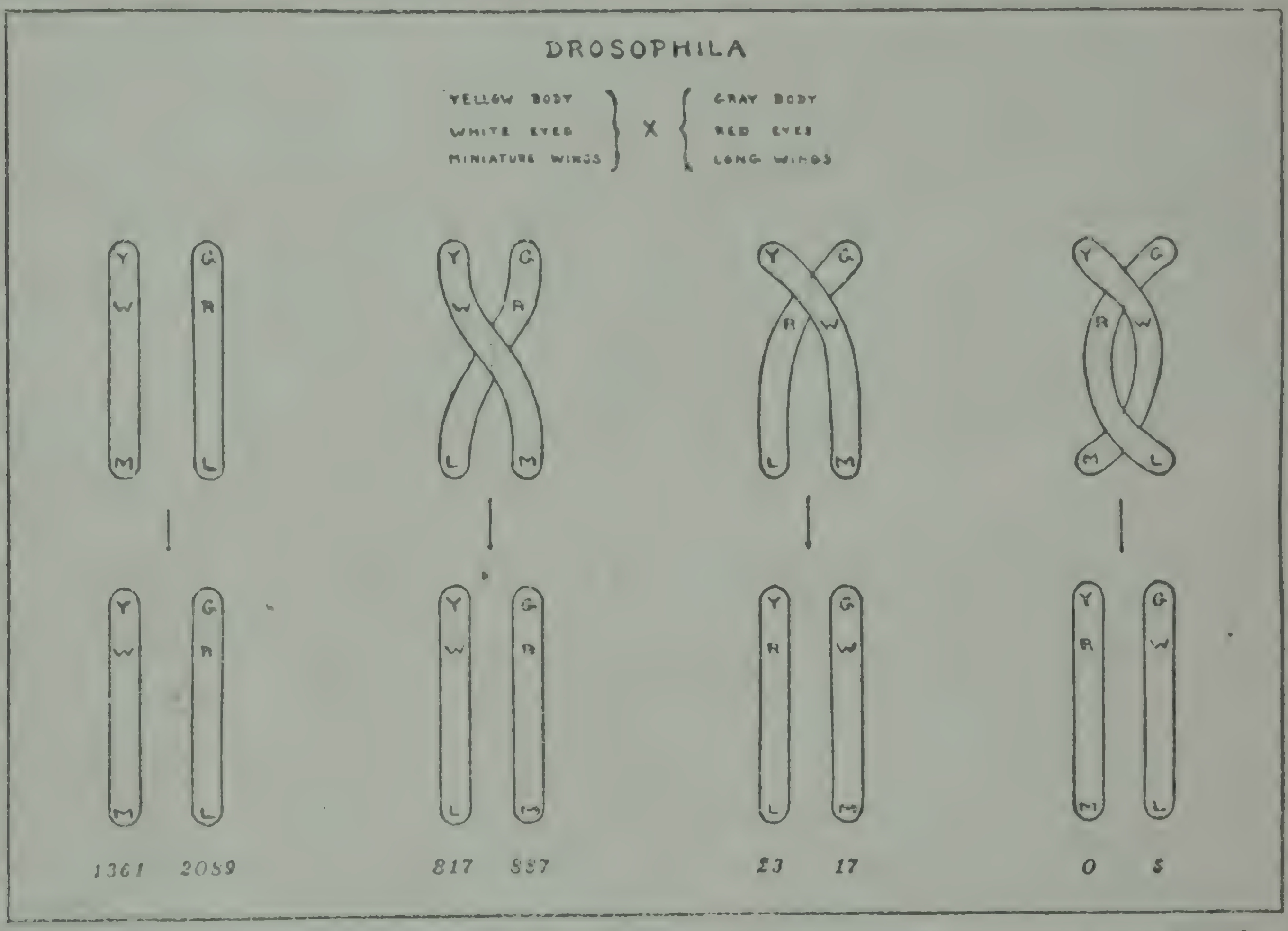


FIG. 693. Diagram to show Formation of New Linkage Groups by Single and Double Crossing-over. (From SHARP, after MORGAN.)

Linkage exists between various other pairs of characters, and the inference is that the genes concerned in the transmission of linked characters must be situated on the same chromosome, and hence be transferred together. The evidence for this, which is again chiefly derived from *Drosophila*, is that there are only as many independent factor-pairs as there are chromosomes in the gamete, or somatic chromosome-pairs. Hence, we conclude that the genes connected with a particular group of characters are located on a particular chromosome. From evidence derived from the detailed study of the phenomenon of crossing-over, it is further believed that for each gene is reserved a particular place in the chromosome, like each bead in a necklace, and that normally this position is faithfully retained in the ensuing generations.

Crossing-over. Although many phenomena of heredity are explicable on the supposition that the genes of a chromosome retain their positions relative to one another from generation to generation, thus giving rise to linkage, yet it appears that this does not always apply as regards the two chromosomes of a somatic pair, carrying therefore similar, though not identical, genes. Let us suppose that, in such a pair, we have two series of genes ABCDEFG, etc., and abcdefg, etc. Then, when reduction division occurs in the formation of the gametes, it may happen that an interchange occurs, so that the gametes may be ABCdefg, etc., and abcDEFG, etc. (Fig. 693).

When this happens, although the linkage persists, there is a certain percentage of departures from the simple Mendelian law of independent association, and it is clear that the further apart on the chromosome any two genes are placed, the more likely will the exchange be to occur.

From the frequency with which this phenomenon of crossing-over happens for particular characters, their relative distance apart on the chromosome has been estimated, and gene-maps of the chromosomes deduced. Double crossing-over also occurs, *e.g.* ABcdeFG, owing to twisting and cleavage having occurred twice in the process of reduction division. These interesting phenomena belong to the special subject of heredity and genetics and cannot be further discussed here.

Mutation. It has been said above that the genic pattern in the chromosomes is reproduced with *almost* unerring accuracy. But not absolutely. All breeders of animals and plants are aware that, occasionally a new feature will suddenly appear and be transmitted thenceforward as an hereditary character. These spontaneously appearing characters are called mutations. On the basis of the gene hypothesis, mutation would be regarded as due to some stable alteration in the properties of the gene concerned, and once the change has occurred in a chromosome of a gamete, a new pattern is established and the mutation persists. We are in complete ignorance as to how these mutations are brought about, so that we cannot, by any means known at present, produce a given mutation, though we can utilise those which occur spontaneously so as to retain the new quality in the breeding of subsequent generations. It has been shown that by the action of X-rays, and by colchicine mutations in *Drosophila* can be made to happen about 150 times as frequently as in nature, and this observation opens up further possibilities of exploring, and perhaps ultimately of controlling, this process. It would seem that the active rays have altered some of the genes, but it appears to be quite fortuitous which ones are affected. It may be that the apparently spontaneous mutations of nature are brought about by the action of cosmic rays acting in a similar manner.

In other instances, mutations have been shown to be associated with a change in the number of chromosomes : this occurs in some plants ; thus Gates has found mutants of *Œnothera* with 15, 20, 21, 22, 23, 27, 28, 29 and 30 chromosomes.

The study of mutation has led to important modifications in our views as to the way in which, in the course of evolution, new forms have appeared. According to the principle of Natural Selection, as put forward by Darwin, any variations, if profitable, would be preserved by natural selection, and thus the process of evolution was a continuous one. The struggle for existence was an essential part of the hypothesis, since by its operation the less well endowed individuals were eliminated, and thus the swamping of the profitable characters was avoided. We now know that mutants arise suddenly, and will breed true, so that swamping will not occur if the mutation is one which brings any biological advantage ; the struggle for existence, though not essential for the production of new and better types, does lead to the loss of those types not able to compete.

For Reference

Crew (1925). Animal Genetics. (London : Oliver and Boyd.)

Crew (1927). The Genetics of Sexuality in Animals. (Cambridge Univ. Press.)

Crew (1933). Sex Determination. (London : Methuen.)

Darlington (1942). The Handling of Chromosomes. (Allen and Unwin.)

Detlejsen (1925). The Inheritance of Acquired Characters. *Physiol. Rev.*, **6**, 244.

Evans, H. M., and Swezy (1929). The Chromosomes of Man, Sex and Somatic. *Mem. Univ. Calif.*, **9**, No. 1, p. 65. (Univ. Calif. Press.)

Hanson (1933). Radiation-Genetics. *Physiol. Rev.*, **13**, 466.

Morgan (1923). The Modern Theory of Genetics and the Problem of Embryonic Development. *Physiol. Rev.*, **3**, 603.

Muller (1947). The Gene. *Proc. R.S.*, **134B**, 1.

Riddle (1931). Factors in the Development of Sex and Secondary Sexual Characters. *Physiol. Rev.*, **11**, 63.

Schrödinger (1944). What is Life ? (Cambridge Univ. Press.)

Stockard (1931). The Physical Basis of Personality. (New York : Norton.)

Strandskov (1944). Physiological Aspects of Human Genetics, Five Human Blood Characteristics. *Physiol. Rev.*, **25**, 445.

Swingle (1926). The Determination of Sex in Animals. *Physiol. Rev.*, **6**, 28.

Wright (1941). The Physiology of the Gene. *Physiol. Rev.*, **21**, 487.

INDEX

Abdominal muscles, respiratory movements of 712
Abdominal reflex, spinal segments concerned in, 215
Abducent nerve, effect of injury to, 376
Abrin, nature of, 701
Abscess, formation of, 699
Absorption spectra, cytochrome showing, **514**
  hæmoglobin derivatives showing, **512**
Acapnia, spinal vasomotor centres affected by, **653**
Acclimatisation, high altitude and factors concerned in, 775, **777**
  hot climates and, 1051
Accommodation, eye showing, 389–391
  time factor of, 192
Acetaldehyde, fat formation from carbohydrate and, 942
Acetate, buffer action of, 65
  fat metabolism and, 945
  oxidation of, **143**
Aceto-acetic acid, fat metabolism and, **945**
  tyrosine metabolism and, 929
Acetone bodies, excretion of, 988
  calorific value of, **791**
  formation of, **945**
Acetylcholine, actions of, 351
  adrenal medulla affected by, 1067
  assay of, **363**
  autonomic nerve stimulation producing, 361
  curare antagonism to, 197
  heart affected by, 588
  motor end-plate and, 197
  pial vessels affected by, 672
  respiratory centre affected by, 765
  salivary glands affected by, 847
  skeletal muscle affected by, 137
  structure of, **29**
  uterine muscle affected by, 165
Ac-globulin, blood clotting and, 530
Achromatic spindle, nature of, 1097
Achromatic threshold, factors affecting, 414
Achroo-dextrin, nature of, 843
Acidæmia, definition of, 504
Acid phosphatase, prostate formation of, 1122
Acid-base balance, renal regulation of, 1011
Acidosis, alveolar carbon dioxide pressure affected by, 757
  ammonia excretion in, 932
Aconitic acid, structure of, **31**
Acromegaly, nature of, **1082**
Acrosome, nature of, 1099
Action potentials, cerebral cortex showing, 319
  definition of, 118
  dorsal roots showing, 247–**248**
  mode of production of, **86**–87
  motoneurone showing, **241**
  nerve showing, **175**
  plain muscle showing, 167–**168**
  pulmonary vagus fibres showing, **760**
  skeletal muscle contraction relation to, **131**
  spinal cord showing, **246, 247**
  sympathetic ganglion showing, **238–239**
Activator, enzyme action and, 93
Actomyosin, nature of, 139
Adaptation, nature of, 4
  sensory receptors showing, **220**
Addison's disease, nature of, 1061
Adenase, action of, 936
Adenine, metabolism of, 936
  structure of, **48, 52**
Adenosine, structure of, 49
Adenosinetriphosphatase. *See* A.T.P.
Adenylic acid, structure of, 49, **52**
Adiadochokinesis, nature of, **341**
A.D.P., structure of, 49
Adrenal cortex, anterior pituitary gland control of, 1063
  ascorbic acid in, 828
  carbohydrate metabolism affected by, 972
  functions of, 1060
Adrenal cortical hormones, actions of, 1061
  rate of production of, 1062
Adrenal gland, anterior pituitary interaction with, 1093
  gonads affected by, 1093
  gonads relation to, 1063
  hair growth affected by, 1032
  history of, 1055
  nervous control of, 1093
  structure of, 1059
  sympathetic nerve supply to, 348
Adrenal medulla, functions of, 1064
  structure of, 1059
Adrenalectomy, anterior pituitary extract injection affected by, **973**
  effects of, 1060
  lactation affected by, 1145
Adrenaline, actions of, 351, **1065**
  adrenal cortex stimulation by, 1062
  anoxia effect on output of, **652**
  anterior pituitary extract effect on reactions to, **970**
  arrectores pilorum affected by, 1028
  blood clotting time affected by, 532
    lactate affected by, **961**
    sugar affected by, **961**
    sugar regulation by, 976
  bronchi affected by, 721
  destruction in body of, 1067
  exercise and effects of, 676
  glomerular filtration affected by, 999
  heart affected by, **590**
    efficiency affected by, 605
    oxygen usage affected by, **605**
  intestinal movements affected by, 893
  liver glycogen affected by, **961**, 963
  nasal mucosa affected by, **709**
  pial vessels affected by, **672**
  portal circulation affected by, 669
  sympathin comparison with, 361
  synaptic response affected by, 239
  temperature regulation aided by, 1038
  tyrosine relation to, 928
  uterine muscle affected by, **164**, 165
  uterus and cervix affected by, 1141
  vasoconstrictor action of, 659
Adrenergic transmission, sites of, **362**
Adrenocorticotropic hormones, actions of, 1086
  adrenal cortex affected by, 1063

Adrenocorticotropic hormones, lactation affected by, 1145
Adsorption, law of, 75
  nature of, 67
Adsorption complexes, factors concerned in formation of, 75
Aerobic oxidases, action of, 99
After-discharge, motoneurones showing, **244,** 245
  reflex response showing, **254–255**
After-image, cause of, 432
  nature of, 426
After-potentials, nature of, **177**
  sympathetic ganglion showing, **238–239**
Age, accommodation of eye change with, 391
  arterial pressure variation with, 613
  basal metabolic rate variation with, 794, **833**
  calorie requirements variation with, **833**
  cardiac output variation with, 568
  protein intake variation with, 834
  pulse rate affected by, 599
  pulse wave velocity change with, 635
Agglutination, blood sedimentation rate affected by, 501
Agglutinin, formation of, 704
Agglutinogens, nature of, 501
Aglomerular fish, renal secretion in, 1007
Agraphia, nature of, 472
Air dust, lungs protection from, 779
Air humidity, heat loss affected by, **1045**
  measurement of, 1047
Air velocity, climatic conditions affected by, **1049**
  measurement of, 1047
Alanine, structure of, 40
Albumins, properties of, 45
Albuminoids, nature of, 51
Albuminuria, causes of, 988, 1002
Alcaptonuria, nature of, 929
Alcohol, calorific value of, **791**
  metabolism of, 839
Alexia, nature of, 472
Alignment chart, blood and, **735**
  body surface calculation by, **794**
Alimentary glycosuria, nature of, 963
Alimentary tract, referred pain from, 353
Alkalæmia, high altitude and production of, 775
Alkali reserve, measurement of, 503
  nature of, 732
Alkalosis, alveolar carbon dioxide pressure affected by, 757
Allantoin, formation of, 937
Allemorph, nature of, 1154
Allergic reaction, intestinal absorption and, 914
Allocortex, nature of, 293
All-or-none law, cardiac muscle and, 545
  nerve showing, 180
  plain muscle showing, 168
  skeletal muscle showing, **127–128**
Alloxan, pancreas affected by, 966
Alpha waves, nature of, 321
Altercursive intubation, nature of, **884**
Altitude, blood volume variation with, 495
Alveolar air, composition of, **724**
Alveolar carbon dioxide pressure, barometric pressure changes effect on, **756**
  constancy of, **725**
  high altitude effect on, **775**
Alveolar gases, partial pressure of, 725
Alveolar membrane, nature of, 709
Alveolar oxygen pressure, high altitude effect on, **775**
Alveoli, structure of, 709
Amacrine cells, retina containing, **403**
Amblyopia, nature of, 376
Amboceptor, nature of, 704
Ametropic eye, accommodation range measurement in, 391
Amine oxidase, ephedrine action on, 1069
Amino-acids, chemical properties of, 39
  deamination of, 100
  electrolyte nature of, 76
  essential and non-essential, **809**
  fate after absorption of, 915
  ketogenic and antiketogenic, 928
  structures of, 40–41
  synthesis in body of, 923
  urine containing, 986
Amino-acid decarboxylases, action of, **97**
Aminophorase, action of, 924
Aminopolypeptidase, action of, 880
Ammonia, excretion of, 932
  kidney formation of, 985
  urea formation from, 926
Ammonium salts, specific dynamic action of, 928
Ammonotelic animal, nature of, 924
Amœboid movement, factors affecting, 169
Amphoteric electrolytes, nature of, 39
Ampulla, nature of, 268
Amylase, action of, 37, 880
  succus entericus containing, 889
Amylopectin, nature of, 37
Amylopsin, action of, 880
Amylose, nature of, 37
Anabolism, definition of, 3
Anacrotic pulse, cause of, 634
Anæmia, cyanosis in, 772
  spleen concerned in, 682
Anæmic anoxia, nature of, 771
Anaerobic bacteria, examples of, 3
Anæsthetics, pupil affected by, 379
  temperature regulation affected by, 1037
Anal sphincters, defæcation and action of, **898**
  innervation of, **896**
Analgesia, nature of, 481
Anaphylactic shock, blood clotting time in, 534
Anaphylaxis, nature of, 51, 706
Anarthria, nature of, 472
Anastomotic circulation, principle governing, **618**
Androgens, capon's comb affected by, **1124**
  nature of, 1123
  prostate affected by, 1122
Androsterone, adrenal cortex containing, 1061
  structure of, 1123
Anelectrotonus, nature of, 189–**190**
Aneurin, actions of, 826
Animal behaviour, cerebral cortex removal affecting, 322
Anisometropia, nature of, 399
Ankle jerk, spinal segments concerned in, 215
Anovulatory cycle, occurrence of, 1107
Anoxæmia, coronary vessels affected by, 602
  nature of, 771
Anoxia, carotid body affected by, 763
  cerebral circulation affected by, 672
  cyanosis in, 772
  respiration affected by, **757**
  types of, 771
  vasomotor centre affected by, 650
Anoxic anoxia, nature of, 771

Anserine, nature of, 140
Anterior pituitary gland, adrenal cortex control by, 1063
  carbohydrate metabolism control by, 969
  correlating action of, 1092
  hormones secreted by, 1080
  hypothalamus relation to, 358
  mammary glands affected by, 1144
  ovary affected by, **1112**
  spermatogenesis affected by, **1121**
  testis affected by, 1125
Anterior pituitary hormones, correlating action of, 1092
Anti-bodies, formation of, 704
Anti-diuretic hormone, action of, 1013, 1088
  factors controlling output of, 1090
Antidromic shock, motoneurone affected by, **241**
  sympathetic ganglion affected by, 239
Antidromic vasodilator nerves, 647
Antigens, nature of, 703
Anti-hormones, nature of, 1058
Antiketogenic amino-acids, nature of, 928
Antiperistalsis, ascending colon showing, 897
  duodenum showing, 874, 893
Antitoxins, production of, 701
Aortic body, function of, 594
  respiratory centre affected by, 764
Aortic nerve, function of, 593
Apatite, nature of, 815
Apex beat, cause of, 562
Aphagocyte, nature of, 697
Aphasia, types of, 471
Apneusis, mechanism of, **762**
  nature of, **752**
Apnœa, cause of, 767
  nature of, 755
  over-ventilation causing, **770**
Apocrine sweat glands, nature of, 1029
Apoplexy, nature of, 315
Aqueous humour, secretion of, **382**
Arachnoid, nature of, 693
Arachnoid villi, nature of, **694**
Arcuate fibres, nature of, 335
Arginase, action of, 926
  corticosterone action on, 1064
Arginine, creatine formation and, 933
  structure of, 41
  urea formation and, 925
Argininephosphate, nature of, 139
Argyll-Robertson pupil, nature of, 289
Arrectores pilorum, action of, 1028
  nerve supply to, 347
Arterenol, nature of, 361
Arterial blood, factors affecting rate of reduction of, **743**
  gas pressures in, **740**
Arterial pressure, adrenaline action on, 659, **1065**
  anoxia effect on, **650**, **652**
  aortic nerve stimulation effect on, **592**
  asphyxia effect on, **650**
  cardiac output effect on, 621
  carotid sinus pressure effect on, **596**
  cerebral circulation affected by, 671
  coronary blood flow affected by, 601
  diuresis caused by rise of, 1012
  exercise effect, **677**
  gravity effect on, 683
  heart volume affected by, **577**
  heart work affected by, 575, **604**
  magnitude of, **538**
Arterial pressure, measurement of, **610**
    in man of, **612**
  plethora effect on, **679**
  post-pituitary extract effect on, 1087
  pulmonary arterial pressure relation to, 665
  pulse wave velocity affected by, 631
  respiration affected by, **766**
    effect on, 667
  spinal cord section effect on, 642
  urine composition affected by, **998**
  urine flow and composition affected by, **1001**
  vagus nerve stimulation effect on, 586
Arterial pulse, characteristics of, 629
Arterial velocity pulse, nature of, 627
Arteries, blood pressure in, 609
  capacity changes with pressure changes in, **539**
  structure of, **539**
Arterioles, chemical regulation of, **658**
  innervation of, 642
  nerve supply to, 347
  pituitrin action on, 1087
  resistance in, 617
  structure of, 539
Arterio-venous anastomoses, sites of occurrence of, 540
  skin containing, 1028
    temperature control by, 1044
Artificial parthenogenesis, methods of, 1101
Artificial respiration, methods of, 718
Arytenoid muscle, action of, 466
Ascorbic acid, actions of, 828
  adrenal cortical hormones relation to, 1062
Aspartic acid, structure of, 41
Asphyxia, nerve affected by, 187
  respiratory movements in, 754, 769
  spleen affected by, 682
  vasomotor centre affected by, **650**
Aspirin, action of, 1053
Astasia, cerebellar ablation causing, 339
Asthenia, cerebellar ablation causing, 339
  cerebellar injury in man causing, **341**
Astigmatism, types of, **398**
Asynergia, nature of, 341
Ataxia, cerebellar lesions causing, 339, 341
  muscle afferent impulses and, 316
Athlete, pulse rate in, 599
  vagal tone in, 588
Athletic records, limiting factors in, **154**, 155
Atmospheric pressure, effects of high and low, 773
A.T.P., acetylcholine synthesis and, 364
  adrenal cortex affected by, 1062
  energy in, 68
  energy-rich bonds in, 52
  myosin and, 148
  skeletal muscle contraction and, 148
    heat production and changes in, 159
  structure of, 49
Atropine, acetylcholine actions affected by, 362
  action of, 352
  eye affected by, 391
Atwater-Benedict calorimeter, use of, **792**
Audito-psychic area, functions of, 303
Auditory centre, location of, 303
Auditory nerve, nerve impulse conduction in, 460
Auditory organ, diagram of, **450**
Auditory reflexes, nervous pathways of, 286
Auditory sensation, central nervous pathways of, **286**
Auerbach's plexus, composition of, 163

Auerbach's plexus, gastric movements affected by, 871
  peristaltic movement and, 891
Auricle, pressure changes in, **558**
  structure of, **549**
  vagus action on, 586
Auricular flutter, nature of, 555
Auriculo-ventricular bundle, course of, **551**
  vagus action on, 587
Auriculo-ventricular node, structure of, 552
Auriculo-ventricular valves, structure of, **549**
Autocoid, nature of, 1057
Autonomic functions, integration of, 354
Autonomic ganglia, functions of, 348
Autonomic nervous system, afferent fibres of, **352**
  anatomy of, **344**
  nature of, 212
  plain muscle innervation by, 163
Autonomic reflex arc, vagus function in, 354
Aviation, circulatory problems in, 684
  oxygen requirements in, 776
Avidin, nature of, 827
Axilla, temperature of, 1035
Axis cylinder, structure of, 172
Axon reflex, hypogastric reflex as, **349**
  metabolites stimulation of, 659
  vasodilator fibres in, **649**

Babinski sign, sleep and, 330
  spinal cord section and, 234
Bacteria, large intestine containing, 894
  leucocyte action on, 699
  serum action on, 700
Baillarger, lines of, **295**
Bainbridge reflex, nature of, 597
Ballistocardiogram, use of, 569
Barány's test, nature of, 272
  cerebellar lesions and, 341
Barger's method for osmotic pressure measurement, **61**
Barometric pressure, alveolar carbon dioxide pressure affected by changes in, **756**
  effects of high and low, 773
Basal ganglia, functions of 313–**314**
  internal capsule relation to, **309**
Basal metabolic rate, age effect on, **833**
  hypophysectomy effect on, 1085
  sex differences in, 1038
  thyroid gland effect on, 1070
Basal metabolism, nature of, 793
Basilar membrane, nature of, **454**
Basophile cells, characters of, **524**
Beats, nature of, 461
Bell and Magendie law, nature of, 181
Benadryl, action of, 662, 706
Benedict's respiration apparatus, use of, **785**
Berger rhythm, factors affecting, **320**
Beri-beri, nature of, 825
β-hydroxybutyric acid, heart utilisation of, 948
β-oxidation, nature of, 944
β-oxybutyric acid, calorific value of, **791**
Beta waves, nature of, 321
Betz cells, position of, 297
Bicarbonate, buffer action of, 65
  carbon dioxide carriage in blood by, 731
  carbonic anhydrase action on, 736
Bidder's ganglion, site of, 542
Bidwell's experiment, nature of, 430
Bile, composition of, 882
Bile pigments, entero-hepatic circulation of, 886
  nature of, 883
Bile salts, composition of, 883
  entero-hepatic circulation of, 886
  fat absorption affected by, 909
  hydrotropic action of, **909**
Bile secretion, 883
  secretin action on, 877
Bilirubin, formation of, 521
  nature of, 883
Biliverdin, formation of, **522**
  nature of, 883
Binocular rivalry, nature of, 433
Binocular vision, advantages of, 445
Bioelectric potentials, nature of, 85
Biological value of proteins, determination of, 809
Biotin, nature of, **827**
Bird, accommodation in eye of, 388
  blood, coagubility in, 531
  eye accommodation in, 447
  red blood cell formation in, **519**
  visual field of, **410**
Birth, circulatory changes at, 1140
Bismuth, excretion of, 895
"**Black-out,**" cause of, 684
"**Black-tongue,**" nature of, 828
Bladder, denervation effect on emptying of, **1023**
  innervation of, **1018**
  pressure measurement in, **1020**
  structure of, 1016
  sympathetic stimulation effect on, 348
  tonus in, 167
Blending, heredity and, 1155
Blindness, cortical damage causing, 305
  nervous lesions causing, 289
  types of, 434
Blind spot, colour fields limitation round, 415, **416**
  measurement of, **408**
Blind zone, extent of, **410**
Blink reflex, reticular formation stimulation effect on, **312**
"**Block,**" spinal cord showing, 263
Blood, alignment chart of, **735**
  alkali reserve of. *See* Alkali reserve.
  amino-acids, pituitary hormones action on, 1085
    protein meal affecting, 915
  calcium, parathyroidectomy effect on, 1076
  carbon dioxide carriage by, **731**
  carbon dioxide dissociation curve of, **732**
  cells. *See* Erythrocyte, Leucocyte, etc.
  cholesterol, fat absorption effect on, 911
  coagulation of, 527–535
    nature of, 493
  composition of, **505**
  electrical conductivity of, 504
  fat, meal taking effect on, 951
  flow, factors controlling intermittent to constant, 622
    measurement, of, 627
    plethysmograph measurement of, **644**
    temperature effect on, 1044
  gases, arterial and venous, **726**
    measurement of, **497–499**
    solubility of, **727**
  gas tensions in, 739
  glycolysis in circulating, 150
    lungs effect on, 605

- **Blood,** groups, types of, 501
  - high altitude effect on, **777**
  - histamine, eosinophile cells and, 524
  - history of, 488
  - lactate, adrenaline action on, **961**
    - factors concerned in, 150
  - laking of, 509
  - lipides of, 911
  - nomogram of, 734
  - osmotic pressure of, 504
  - oxygen content of, 500
  - oxygen dissociation curve of, 727
    - pregnancy effect on, **1138**
  - oxygen saturation, anoxia effect on, **772**
    - coronary blood flow affected by, **602**
    - pregnancy effect on, **1135**
  - oxygen uptake by, 729
    - in lungs by, 730
  - oxygenation, carbon dioxide dissociation curve affected by, **734**
  - pH of, 504
    - renal regulation of, 1011
  - phosphate, parathyroidectomy effect on, 1076
  - pigments, synthesis of, 516
  - plasma. *See* Plasma.
  - platelets. *See* Platelets.
  - pressure, 609
    - salivary duct pressure relation to, 848
    - sleep effect on, 330
    - *See also* Arterial pressure, Venous pressure, etc.
  - reaction of, 503
  - red cells. *See* Erythrocytes.
  - respiratory function of, 726
  - sedimentation rate of, 501
  - serum. *See* Serum.
  - specific gravity of, 501
  - spleen action on, 682
  - sugar, adrenal cortical hormone effect on, **973**
    - adrenaline action on, **961**
    - factors regulating concentration of, 1091
    - glucose ingestion effect on, **953**
    - hunger relation to, 900
    - insulin action on, 968
    - lactation affected by, 1146
    - liver glycogen relation to, 960
    - normal regulation of, 975
    - regulation of, 954
    - stomach movements affected by, 874
  - transfusion of, 503
  - velocity of, 622
    - measurement of, 624
  - vessels, branching effect on flow in, **618**
    - nervous control of, 642
    - peripheral tone of, 654
  - viscosity of, 501
    - corpuscular content affecting, **619**
  - volume, measurement of, 494
    - shock relation to, 663
    - sleep effect on, 330
    - temperature effect on, 1046
  - white cells. *See* Leucocytes.
- **Body** fluid, temperature effect on movement of, 1046
- **Body** surface, metabolic rate relation to, 793
  - nomogram for calculation of, **794**
- **Body** temperature, limits to variation in, 1051
  - regulation of, 1036
    - history of, 981
  - shivering relation to, 1054
- **Body** temperature, sympathetic system removal effect on, 350
- **Body** weight, basal metabolic rate affected by loss of, 795
- **Bohr** effect, nature of, 729
- **Boiler-maker's** disease, basilar membrane damage in, 457
- **Bone,** composition of, 815
  - growth of, 815
  - parathyroid effect on, 1078
  - vitamin A effect on growth of, **821**
  - vitamin C effect on growth of, 830
  - vitamin D effect on growth of, 823
- **Bone,** marrow, blood cell formation in, 518
  - granulocyte formation in, 525
- **Borate,** buffer action of, 65
- **Bowman's** capsule, blood supply to, 991
  - fluid collection from, **996**
- **Bragg-Paul** pulsator, use of, 719
- **Brain,** blood flow to, 629
  - blood supply to, 670
  - glucose usage by, 670
    - and pyruvate usage by, 746
  - meninges of, 693
  - oxygen usage of, 670
  - pressure in, 671
  - section, result at different levels of, **273**
  - starvation and loss in weight of, **797**
  - stem, structure of, **210, 211**
  - sympathetic nerve supply to, 347
  - vitamin B_1 action on, 826
- **Bread,** gastric secretion response to, 869
  - war-time modification of, 837
- **Breath-holding** tests, nature of, 768
- **Breath** sounds, nature of, 714
- **Brewster's** stereoscope, use of, **448**
- **Brightness** difference threshold, light intensity effect on, **417**
- **Broca's** area, function of, 471
  - location of, 300
- **Bronchi,** air dust removal by, 779
  - movements of, 722
  - parasympathetic stimulation effect on, 350
  - respiratory movements of, 712
  - structure of, 708
- **Bronchial** murmur, nature of, 714
- **Bronchial** musculature, factors affecting, 720
- **Bronchioles,** structure of, 709
- **Brown-Séquard** syndrome, nature of, 235
- **Brunners'** glands, secretion by, 888
- **Bubble** stromuhr, use of, **624**
- **Buffer** mixtures, pH range of, **65**
- **Buffer** systems, action of, **64**
- **Buffy** coat, nature of, 501
- **Bulbo-cavernosus-muscle,** action of, 1018
- **Bulbo-spinal** preparation, nature of, 232
- **Burch's** experiment, nature of, **431**
- **Butter,** formation of, 1148

- **Cadaverine,** nature of, 97
- **Caffeine,** diuretic action of, 1015
  - structure of, 48
- **Caisson** disease, nature of, 778
- **Calcarine** cortex, retinal representation in, **304**
- **Calciferol,** nature of, **823**
- **Calcification,** factors concerned in, 816
- **Calcium,** absorption from intestine affected by, 907
  - balance sheet of, **812**
  - blood clotting affected by, 528
  - bone formation and, 816

- **Calcium**, cardiac muscle affected by, **581**
  - cell membrane permeability affected by, 16
  - excretion of, 985
  - ganglion cells affected by, 366
  - histamine action affected by, 721
  - metabolism, parathyroid effect on, 1078
    - vitamin D action on, 823
  - milk clotting and, 862
  - nerve accommodation affected by, 193
  - nerve affected by, 186
    - response to stimulus affected by, 175
  - trypsinogen activation by, 879
- **Calcium** absorption, factors affecting, 813
- **Calorie**, definition of, 789
- **Calorie** requirements, factors affecting, 832
- **Calorimetry**, history of, 981
  - indirect, 792
- **Cancer**, growth of, 1153
- **Cane** sugar, gastric juice action on, 861
- **Capillaries**, blood pressure in, 609
  - capacity of, 540
  - circulation through, 636
  - contractility of, 637
  - exercise effect on muscle, **742**
  - fragility of, 638
  - innervation of, 661
  - leucocyte emigration from, **698**
  - nerve supply to, 347
  - permeability of, 81
  - resistance to blood flow offered by, 617
  - structure of, 539
- **Capillary** electrometer, principle of, 118
- **Capillary** permeability, factors affecting, 689
  - shock relation to, 663
  - tissue fluid formation and, 687
- **Capillary** pressure, factors affecting, 638
  - magnitude of, **538**
  - measurement in man of, 638
  - tissue fluid formation affected by, 688
- **Capillary** tone, pituitrin action on, 1087
- **Capon's** comb, androgens and œstrogens effect on, **1124**
- **Carbamino-acids**, nature of, 39
- **Carbamino-compounds**, hæmoglobin as, 737
- **Carbohydrate**, anti-ketogenic action of, 947
  - calorific value of, **791**
  - chemical nature of, 30
  - exercise and usage of, 154
  - fat as source of, 944
    - formation from, 941
  - gastric digestion of, 862
  - intermediary metabolism of, **962**
  - intestinal absorption of, 913
  - ketosis affected by, 965
  - pancreatic digestion of, 880
  - salivary digestion of, 843
  - sparing action on protein of, 802
  - specific dynamic action of, 803
  - stomach emptying time affected by, 873
- **Carbohydrate** metabolism, 952
  - bilirubin formation affected by, 522
  - exercise effect on, 806
  - hormonal control of, 969
  - insulin action on, 969
  - lactation affected by state of, 1145
  - respiratory quotient showing amount of, **805**
  - starvation effect on, 797
  - vitamin B_1 action in, 826
- **Carbon**, balance sheet of, 788
- **Carbon** dioxide, absorption coefficient of, **726**
  - blood carriage of, **731**
- **Carbon** dioxide, blood carriage of, velocity affected by, **625**, 627
  - carotid body affected by, 763
  - cerebral circulation affected by, 672
  - dissociation curves, blood and plasma comparison, **732**
    - oxygen effect on, **734**
  - Fick principle use of, 569
  - glycogen formation from, 959
  - heart affected by, 591
  - oxygen dissociation curve affected by, **728**
  - pressure, air and blood comparison, **740**
  - reduction rate of arterial blood affected by, **743**
  - respiration affected by, **756**
    - after vagotomy affected by, **761**
  - skin loss of, 1032
  - tissue production of, 745
  - vasomotor centre affected by, **651**
- **Carbonic** anhydrase, action of, 97, 736
  - oxyntic cells and, 863
- **Carbon** monoxide, absorption coefficient of, **726**
  - blood volume measurement by use of, 494
  - hæmoglobin determination by use of, 497
  - toxic effects of, 778
- **Carboxyhæmoglobin**, dissociation curve of, 779
  - formation of, 511
  - molecular weight of, 45
  - oxygen dissociation curve affected by, 729
- **Carboxylase**, action of, **97**
- **Carboxypolypeptidase**, action of, 880
- **Cardiac** cycle, coronary flow variations in, **603**
  - pulse tracings relation to, 633
  - sequence of events in, 556–561
- **Cardiac** efficiency, adrenaline effect on, 605
- **Cardiac** impulse, cause of, 562
- **Cardiac** index, nature of, 568
- **Cardiac** murmurs, cause of, 563
- **Cardiac** muscle, chronaxie of, 114
  - ions action on, **581**
  - length effect on response of, 577
  - metabolism of, 604
  - myoglobin function in, 744
  - properties of, 544
  - structure of, **541**
- **Cardiac** nerves, action of, 588
- **Cardiac** output, arterial pressure affected by, 621
  - basal metabolic rate calculation from, 795
  - circulation time relation to, 623
  - coronary vessel tone affected by, 603
  - determination of, 568
  - exercise effect on, 675
  - factors affecting, 573
  - heart work affected by, **604**
- **Cardiac** reflexes, nature of, 595
- **Cardiac** sphincter, deglutition and action of, **854**, 857
- **Cardiometer**, use of, 568
- **Cardio-accelerator** centre, location of, 592
- **Cardio-aortic** nerve, function of, 593
- **Cardio-inhibitory** centre, action of, 592
  - respiratory centre action on, **600**
  - vagus nerve excitability affected by state of, 587
- **Cardio-pneumatic** movements, nature of, 570
- **Carnivor**, protein storage in, **801**
- **Carnosine**, nature of, 140
- **Carotene**, vitamin A relation to, 822
- **Carotid** body, respiratory centre affected by, **763**
  - structure of, 595

Carotid sinus, innervation of, **594**
  nerve endings in, **218**
  perfusion of, **596**
  reflex secretion of adrenaline from action of, 1067
  respiratory centre affected by, **764, 766**
  stretch receptors in, 594
Carotid sinus nerve, function of, 354
  impulses in, **597**
Casein, iso-electric pH of, 74
Caseinogen, amino-acids in, 44
  gastric juice action on, 862
  mammary gland manufacture of, 1147
  properties of, 47, 1149
Castration, effects of, 1104
  pituitary changes following, 1079
  seminal vesicles affected by, **1125**
Catalase, action of, 97
Catalyst, definition of, 88
Cataract, treatment of, 388
Catelectrotonus, nature of, 189–**190**
Cathepsin, action of, 915
  nature of, 861
Cathode ray oscillograph, use of, **118,** 119
Cell, differentiation of, 19–21
  electric potentials in, 85–87
  plasma membrane of, 15
  structure of, 9
    phospholipides in, 951
Cell count, blood viscosity affected by, **619**
Cell membrane, permeability of, 15, 81–82
  permeability in muscle of, 137–138
  structure of, 16, **17**
Cellulose, fæces composition affected by dietary, **900**
  properties of, 38
Central excitatory state, nature of, 239
  time relations of, **258**
Central inhibitory state, nature of, 239
  time relations of, **259**
Central nervous system, acetylcholine formation in, 364, **365**
  functional levels in, 229
  history of, 198
  hypoglycæmia effect on, 956
  nerve fibre structure in, 204
  structure of, 209, **210**
  theory of inhibition in, **260**
  vitamin A action on, 821
Central reflex time, factors controlling, 253
Centrosome, cell division affecting, 1096
  function of, 13
Cephalin, thromboplastin containing, 531
Cerebellar nuclei, connections of, 334
  stimulation of, 340
Cerebello-frontal tract, course of, 338
Cerebello-rubro-thalamic tract, course of, 338
Cerebello-tegmental tracts, course of, 337
Cerebellum, connections of, 335, **336**
  functions of, 338
  structure of, **333**
Cerebral circulation, 670–672
Cerebral cortex, animal behaviour affected by removal of, 322
  associative functions of, **301**
  autonomic functions control by, 359
  cerebellar connections with, **336, 337**
    stimulation effect on, 339
  conditioned reflexes after damage to, 328
  cytoarchitectonic map of human, **294**
  electrical activity of, 319
  eye movement representation in, 376
Cerebral cortex, functions of, 318
  hypothalamic connections with, 356, 358
  localisation of function in, **299**
  micturition control by, 1025
  motor areas of, 305
  neurone types in, **296, 297**
  removal, effects of, 233
  sensory areas in, 302
  speech centres in, 471
  stimulation, respiration affected by, 753
  structure of, 293
  thalamus connections with 291
Cerebral hemispheres, development of, 293
Cerebro-spinal fluid, circulation of, 695
  calcium in, 814
  location of, 693
Cerebrum, coronal section of, **314**
Ceruminous glands, secretion of, 1029
Cervix uteri, parturition effect on, 1140
  reactions of, 1141
Charpentier's bands, nature of, **426**
Cheese, formation of, 1149
Chemotaxis, leucocyte action and, 699
  nature of, 3, 483
Cheyne-Stokes breathing, mechanism of, 771
Chloride excretion, post-pituitary regulation of, 1013
Chloride, daily requirement of, **812**
  ptyalin action and, 843
  reabsorption in kidney of, **1005**
  skin storage of, 1032
  sweat containing, 1029
  urine containing, 984
Chlorocruorin, nature of, 516
Chlorophyll corpuscles, starch formation by, 22
Cholane ring, adrenal cortical hormones and, **1061**
  structure of, **28**
Cholanic acid, nature of, 883
Cholecystokinin, action of, 886
Choleglobin, nature of, 521
Choleic acids, nature of, 883
Cholesterol, adrenal cortical hormones relation to, 1062
  bile salts relation to, 883
  intestinal absorption of, 913
  metabolism of, 952
  structure of, **28**
  vitamin D relation to, **824**
Cholic acid, bile secretion affected by, 884
  structure of, **28**
Choline, acetylcholine synthesis and, 364
  lipotropic action of, 949
  structure of, **29**
Choline acetylase, action of, 364
Choline esterases, action of, 362
  motor end plate containing, 197
Cholinergic nerves, sweat glands supplied by, 1031
Cholinergic transmission, sites of, **362, 365**
Chondroitin, nature of, 50
Chorda tympani, effect of stimulation of, 847
Chorionepithelioma, hormone excretion with, 1134
Chorionic gonadotropin, nature of, 1132
Chorionic villi, formation of, **1132**
Choroid plexus, cerebro-spinal fluid formation by, 694
Chromaffine substance, nature of, 1060
Chromatic aberration, eye correction of, 444
  eye showing, **393**
Chromatin, composition of, 17, 48

Chromatography, nature of, 75
Chromatolysis, nature of, 204
Chromophobe cells, nature of, 1079
Chromoproteins, nature of, 47
Chromosomes, cell division affecting, 1096
  gene maps of, 1159
  heredity and, 1153
  species number of, 1097
Chronaxie, muscles showing different values of, 194
  nature of, 114
Chyle, nature of, 686
Chylomicrons, fat meal affecting, **910**
  nature of, 506
Chymosin, action of, 862
  caseinogen affected by, 1149
Chymotrypsin, action of, 880
Ciliary movement, nature of, 169–170
Ciliary muscle, action of, 389
  nerve supply to, 350, 390
Circulation, artificial schema of, **620**
  capacity changes in, 621
  history of, 489
  hot climate effect on, 1051
  hydrodynamics of, **615**
Circulation rate, reduced barometric pressure effect on, 776
Circulation time, measurement of, 623
Circulatory shock, events in, 663
Circus movement, nature of, **555**
Citrate, acetylcholine synthesis and, 364
  blood clotting affected by, 528
  semen containing, 1123
Citrin, action of, 830
Citrulline, structure of, 41
  urea formation and, **927**
" **Clasp-knife** " reaction, decerebrate preparation showing, 265
Clearance tests, nature of, 1002
Climacteric, nature of, 1103
Climate, factors affecting tolerability of, **1050**
Clonus, nature of, **129–130**
" **Closing** tetanus," nature of, 189
Clothing, temperature regulation affected by, 1046
Cluster hypothesis of colour vision, 439
Cobalt, hæmopoietic factor and, 520
  function in diet of, 813
Cocaine, adrenaline action affected by, 351
  taste affected by, 484
Co-carboxylase, structure of, **94**
Cochlea, structure of, 453
Cochlear telephonic action, 456
Cochlear microphonic action, 455
Coefficient of utilisation, exercise effect on, 676
  factors affecting, 742
Co-enzyme I, E_h of, 101
  lactic dehydrogenase action on, **98**
  structure of, **93**
Co-enzyme system, nature of, 93
Coitus, events in, 1127
  nervous mechanism of, 1128
Colchicine, cell division affected by, 1097
  mutations induced by, 1159
Cold, metabolic rate affected by, 1037
Cold spots, distribution of, **475**
Collagen, gastric digestion of, 862
  properties of, 51
Colloid goitre, nature of, **1074**
Colloid osmotic pressure, measurement of, **71**
  nature of, 57
Colloidal ions, nature 76
Colloids, coagulation of, 77
  electrical properties of, 73
  interaction between, 77
  properties of, 69
Colon, defæcation and action of, **898**
Colostrum, nature of, 1146
Colour, difference threshold of, **423**
  extinction curve of, **415**
  intensity threshold for, 414
  threshold, limits of, 422
  visual fields of, **410,** 415
Colour blindness, detection of, 434
  sex-linked inheritance of, 1158
  types of, 436
Colour index, nature of, 497
Colour patch apparatus, **424,** 425
Colour triangle, nature of, **373**
Colour vision, lateral geniculate body structure and, 288
  theories of, 437
Combination tones, nature of, 451, 462
Comfort, physical factors affecting, 1051
Commutator, neurone activity and, 207, **208**
Compensatory pause, nature of, **547**
Complement, nature of, 704
Complemental air, value of, **716**
Complementary colours, nature of, 372
Complement-fixation test, 705
Compressor urethra, action of, 1018
Concealed reflex, nature of, 266
Conchiolin, nature of, 51
Conditioned reflex, formation of, 324
  gastric secretion and, **865**
  laws of, 325
  nature of, 323
Conditioned stimulus, nature of, 324
Conduction, heat loss by, 1039
Cone, diameter at fovea of, **419**
  diameter, visual acuity and, 416
  fovea centralis containing, 408
  light affecting shape of, 405
  structure of, **404**
  theories of vision and, 437
Conjunctival reflex, nature of, 380
Consciousness, site of, 318
Consonance, nature of, 462
Consonants, production of, 470
Constrictor pupillæ, nerve supply to, 350
Contact stimuli, righting reflexes and, **281**
Contraction remainder, nature of, 135
Contracture, types of, 136
Contrast phenomena, eye showing, 430
Convection, heat loss by, 1039
Convergence, neurone activity showing, **208**
  reflex response and principle of, **256**
Copper, function in diet of, 813
  functions of, 23
  hæmoglobin formation and, 520
Coproporphyrin, formation of, 523
  nature of, 516
" **Core-conductor** " model, nature of, **192**
" **Cori cycle,**" nature of, **961**
" **Cori ester,**" structure of, **34**
Cornea, focal length of, 387
  nerve endings in, **217**
  radius of, **386**
  sensation in, 481
  structure of, 377
Coronary circulation, arterial pressure effect on, 601
  blood oxygen saturation effect on, **602**
  cardiac cycle effect on, **603**

Coronary circulation, measurement of, 573
  volume of, 600
Coronary vessels, adrenaline action on, 591
  nervous control of, 602
  vasodilator nerves to, 646
Corpus Arantii, nature of, 550
Corpus luteum, formation of, **1107**
  mammary glands affected by, 1104
  menstrual cycle and changes in, 1118
  œstrus cycle affected by, 1113
Corpus striatum, functions of, 313, **314**
  motor cortex connections with, 311
Corti, organ of, **455**
Cortical epilepsy, nature of, 308
Cortical unit, nature of, **298**
Cortico-cerebellar tracts, destination of, 337
Cortico-nuclear tract, nature of, 311
Cortico-spinal tracts, composition of, 226
  spinal shock and, 230
Corticosterone, lactation affected by, 1145
  structure of, **28, 1061**
Cough, reflex mechanism of, 759
Coupled reactions, nature of, 68
Cranial autonomic afferent fibres, course of, 354
Crayfish adductor, double innervation of, **162**
Creatine, metabolism of, 933
  structure of, **53**
Creatinephosphate, adrenal cortex affected by, 1062
  properties of, 139
  skeletal muscle contraction and, 149
    heat production and changes in, 159
  structure of, **53**
Creatinine, excretion of, 986
  excretion, diet effect on, **924**
  metabolism of, 933
Creatinine clearance, value of, 1003
Creatinuria, cause of, 934
Crescent cells, nature of, 844
Cretin, nature of, 1069
Crico-arytenoid muscles, action of, 466
Crico-thyroid muscle, action of, 466
Crista acoustica, nature of, 268, **269**
Crossed extensor reflex. *See* Extensor reflex.
Crossing-over, heredity and, 1158
Crotonic aldehyde, fat formation from carbohydrate and, 942
Cryptopyrrole, nature of, **515**
Crystallin, properties of, 46
Crystalline lens. *See* Lens.
C.U., nature of, 1133
Cupola, nature of, 268, **269**
Curare, acetylcholine antagonism to, 197
  action of, 351
  end-plate potentials affected by, 196
  eserine antagonism to, 188
  sympathetic ganglia affected by, 240
  temperature regulation affected by, 1037
Cutaneous receptors, types of, 216, **217**
Cutaneous sensations, 474
Cutis, structure of, 1028
Cyanhæmoglobin, nature of, 513
Cyanide, circulation time measurement by, 623
  cytochrome affected by, 514
  cytochrome oxidase affected by, 99
  heart metabolism affected by, 747
  vasomotor centre affected by, **652**
Cyanosis, nature of, 772
Cyclol structure, nature of, 43
Cysteine, acetylcholine synthesis and, 364
  lactalbumin containing, 1151
  metabolism of, 931
  structure of, 40
Cystine, lens containing, 385
  structure of, 40
Cytochrome, heart containing, 747
  location of, 13
  nature of, 514
  skeletal muscle containing, 140
Cytochrome oxidase, action of, **99**
Cytochrome peroxidase, action of, 100
Cytochrome reductase, action of, 100
Cytolysins, formation of, 704
Cytoplasm, fluid nature of, **14**–15
  structure of, **12, 13**
Cytosine, structure of, **49**

Dark adaptation, nature of, 414
Dead space, measurement of, 725
  value of, 718
Deafness, basilor membrane damage in, 457
Deamination, enzymes concerned in, 100
Decapitate preparation, nature of, 229
Decerebrate preparation, lengthening and shortening reactions in, 265
  properties of, **232**
Decerebrate rigidity, cerebellar stimulation effect on, 339
  mechanism of, **273–274**
  nature of, 229, **232**
Decerebration, blood sugar affected by, 964
Decibel, definition of, 464
Decidua, formation of, 1131
Decidual cells, nature of, 1131
Decorticate preparation, properties of, 233
Defæcation, mechanism of, 897
  spinal cord section affecting, 234
Defence, chemical methods of, 700
Defibrinated blood, nature of, 493
D.F.P., action of, 351, 362
Deglutition, nervous mechanism of, 858
  process of, 852
Deglutition apnœa, mechanism of, 767
Dehydration, effects of, 1052
  treatment of, 811
Dehydrogenases, actions of, 96
  heart containing, 747
Deiter's nucleus, connections of, **269,** 270
  decerebrate rigidity and, **273**
  function of, 233
Delayed conditioned reflex, nature of, 326
Delta waves, nature of, 321
Demarcation current, nature of, 116
Denaturation, nature of, 38
Dendrons, nature of, 203
Denervated tissues, adrenaline action on, 1067
Depolarisation, nerve impulse and, **194**
  sense receptor stimulation and, 221
Depot fat, composition of, 939
  functions of, 942
Depressor nerve, function of, 593
Depressor reflex, nature of, 595, **655**
Depth perception, hypotheses of, 448
Dermatographism, nature of, 662
Desoxycorticosterone, actions of, 1062
  kidney affected by, 1010
  structure of, **1061**
Deuterium, amino-acid metabolism studied by, 923
  depôt fat lability shown by, 942
  nature of, 920
Dextrins, properties of, 37–38
Dextrocardiogram, nature of, 566

Diabetes insipidus, lesions producing, 358
nature of, 1013, 1089
thirst in, 901
Diabetes mellitus, alveolar carbon dioxide pressure affected by, 757
anterior pituitary cause of, 1085
causes of, **977**
G : N ratio in, 944
metabolism in, 967
thirst in, 901
treatment of, 978
Diabetogenic hormone, actions of, 971
diet effect on secretion of, 972
growth hormone relation to, 1084
Dialysis, nature of, 55, **56**
Diaphragm, respiratory movements of, 710
vomiting and action of, 874
Diastole, cardiac events in, 556
heart filling in, 567
pressure changes in heart in, **559**
Diastolic murmur, cause of, 563
Diastolic pressure, measurement of, **611**
measurement in man of, 612
Diastolic volume, oxygen usage of heart relation to, 579
Diathermy, principle of, 115
Dibenamine, actions of, 351
Dicarboxylic amino acid, transamination and, 923
Dicrotic wave, nature of, 634
Dicoumarin, action of, 530
vitamin K relation to, **825**
Diet, adequate types of, **838**
calorie requirements of, 832
calorie value, vitamin B_1 requirements related to, 826
class differences in, **836**
fæces composition affected by, **900**
foodstuffs distribution in, **834**
nitrogen excretion variation with, **924**
pancreatic juice composition affected by, **881**
protein variation in, 802
starvation metabolism affected by previous, 799
Difference thresholds, types of, 423
Difference tones, nature of, 451, 462
Differential cell count, average values of, 523
Diffusion coefficient, measurement of, 54
Diffusion, complexities of, 79
Digestion, history of, 781
time course of, 918
Digitalis, heart affected by, 579
Dihydroxyacetone, structure of, **30**
Di-iodotyrosine, structure of, 40, **1072**
Dilator pupillæ, nerve supply to, 347
Dilemma, reaction time and, 319
Dilution diuresis, nature of, 1012
Dinitro-α-naphthol, temperature affected by, 1037
Diodrast T*m*, value of, 1009
Diœstrus, vaginal changes in, **1108**
Dioptre, definition of, 387
Diphasic action potential, cause of, **119**
Diphosphoglyceric acid, structure of, **53**
Diphtheria anti-toxin, production of, 702
Diploid number, nature of, 1097
Diplopia, strabismus causing, 376
Disaccharide, nature of, 36
"**Discobolus position,**" nature of, 272
Discrimination, conditioned reflexes showing, 325
cortical areas concerned in, 302
Discrimination, touch showing, 478
Discus proligerus, hyaluronidase action on, 1128
nature of, 1106
Disease, food intake affecting, 834
Dissonance, nature of, 461
Diuresis, diabetes mellitus causing, 977
hormonal factors controlling, 1090
types of, 1012
Diuretic hormone, action of, 1088
Diuretics, action of, 1015
osmotic type of, 1006
Divergence, neurone activity showing, 207, **208**
Diver's palsy, nature of, 778
Dominant character, nature of, 1154
"**Dominator,**" nature in vision of, 440
Donnan equilibrium, nature of, 82–84
red blood cells and, 510
Donnan ratio, nature of, 735
Dorsal roots, potential changes in, 247, **248**
Dorsal root section, effects of, 228
Dorsal root stimulation, vasodilatation caused by, **648**
Douglas bag, use of, **723**, 787
Draper's law, nature of, 373
Drosophila, sex chromosomes in, **1157**
Ductus endolymphaticus, nature of, 268
Ductus venosus, fœtal circulation in, 1137
Duodenal bulb, stomach emptying and, 871
Duodenal contents, sampling of, 875
Duodenum, food absorption in, **919**
gastric secretion stimulation from, 866
movements of, 893
secretion in, 888
stomach emptying and pressure in, 873
Duplex theory of colour vision, 437
Duplicity theory of vision, 418
Dura mater, nature of, 693
Dwarfism, nature of, **1083**
Dye, blood volume measurement by use of, 494
Dyslysin, nature of, 886
Dysmetria, nature of, 341
Dyspnœa, nature of, 717, 769

Ear, diagram of, **450**
Ear ossicles, movements of, 452
Eck fistula, nature of, 669
Ectoplasm, properties of, 14
Edestin, amino-acids in, 44
iso-electric *p*H of, 74
properties of, 46
Edridge Green lantern, use of, **435**
Effective temperature, nature of, **1048**
Egg albumin, amino-acids in, 44
iso-electric *p*H of, 74
molecular weight of, 72
properties of, 45
Egg white, biotin action and, 827
intestinal absorption of, 914
Egg yolk, vitamin A content of, 822
Ehrlich cells, heparin formation in, 533
Ejaculation, nerve centre of, 1128
Elastin, properties of, 51
Electrical stimuli, types of, 110
Electro-acoustic effects, cause of, 456
Electrocardiogram, recording of, 564
vagus nerve stimulation effect on, 587
Electro-dialysis, nature of, 80
Electroencephalograms, types of, **321**
Electro-endosmosis, nature of, 80

Electrolytes, osmotic pressure of, 61
proteins as, 76
Electroneuronography, nature of, 224
Electrophoresis, nature of, 74
Electrotonic current, nature of, **191, 192**
Electrotonic potential, motor end plate and, 196
nature of, 190–**191**
spinal roots showing, 247, **248**
sympathetic ganglia showing, 240
synaptic transmission and, **237**
Electrotonus, nature of, **189**
Eleidin, skin containing, 1028
Embryo, development of, 1131
nutrition of, 1134
tissue cultures from, 21
Emergency theory of adrenaline action, 1068
Emmetropic eye, accommodation range in, 391
Emotion, heart rate affected by, 598
pupil size affected by, 379
sweating response to, **1030**
Emphysema, nature of, 715
Emulsification, nature of, 27
Emulsin, action of, 89
Emulsoid, definition of, 70
Endocardiac pressure, measurement of, **557**
Endocrine actions, correlation of, 1091
Endocrine grafts, factors controlling growth of, 1092
Endocrine organs, history of, 1055
methods of investigation of, 1058
nervous control of, 1093
End-plate potentials, nature of, 196
Endurance test, nature of, 768
Energy, conservation in body of, **2**
Energy rich bonds, use of, 68
Energy-rich phosphate, skeletal muscle contraction and, 149
Enterocrinin, nature of, 889
Enterogastrone, actions of, 867
Entero-hepatic circulation, nature of, 886
Enterokinase, action of, 93, 879
Environmental constant, nature of, 1040
Enzyme action, kinetics of, **91**
Enzymes, catalytic nature of, 90
classification of, **95**
nature of, 90
specificity of, 89
synthetic action of, 94
Eosinophile cells, character of, **524**
infection and action of, 699
Ephapses, nature of, 237
Ephedrine, action of, 351, 1069
Epicritic sensation, nature of, 481
Epididymis, ciliary movement in, 169
sperm storage in, 1121
structure of, 1119
Epigenesis, theory of, 1152
Epilepsy, cortical stimulation producing, 308
Equiline, nature of, 1111
Equivalent temperature, nature of, 1047
Erection, mechanism of, 1120
nerve centre of, 1128
Erector penis, action of, 1127
Erepsin, action of, 39
nature of, 889
Ergosterol, vitamin D relation to, **823**
Ergotamine, adrenaline actions affected by, 1066
sympathetic system affected by, 647
Ergotoxine, sympathetic system affected by, 351

Erythroblast, nature of, 519
Erythrocytes, average data of, 500
blood viscosity affected by number of, **619**
buffering power of, 733
chemistry of, 510
composition of, **505**
formation, uric acid excretion affected by, 938
fragility of, 509
gases solubility in, **727**
hæmolysin action on, **705**
high altitude effect on, **777**
life history of, 517
number of, 493
osmotic pressure measurement by use of, 59
permeability of, 16, 81
size of, 507
specific gravity of, 501
splenectomy effect on, 682
sugar content of, 953
volume of, 496
Erythro-dextrin, nature of, 843
Eserine, action of, 351, 362
curare antagonised by, 188
eye affected by, 391
mode of action of, 589
Ether, pulmonary circulation time measurement by, 623
Ethereal sulphates, excretion of, 932
Eupathioscope, use of, 1047
Eustachian canal, nature of, 452
Evaporation, heat loss by, 1040
Eve's rocker, use of, 719
Excitability, temperature effect on, 126
time factor in, 111
Excitation, membrane permeability affected by, 87
Exercise, body temperature affected by, 1035
capillary area in muscle with, **742**
cardiac nerves importance in, 584
circulatory changes in, 673–678
creatinine excretion affected by, 934
foodstuffs used in, **943**
lactic acid production in, 150–**151**
lymph flow affected by, 691
mechanical efficiency of, 155
metabolic changes in, 804
oxygen consumption in, 150–**152**
protein requirements in, 806
respiratory changes in, **674**
respiratory quotient during 153–154
second wind in, 769
skeletal muscle capillary area increase in, 638
glycogen changes ni, 960
spleen affected by, 682
sweating induced by, 1031
sympathetic system removal effect on, 350
uric acid excretion affected by, 938
ventilation rate affected by, 768
white cell count in, 524
Exophthalmic goitre, nature of, **1075**
Exotoxins, body reaction to, 701
Expiration, muscles used in, 711
Expiratory centre, location of, 750
Expired air, composition of, 723
measurement of, 722
Expression, function of, 472
Extensor muscles, chronaxie of, 194
Extensor reflex, ipsilateral, 266
nature of, **230,** 231
nociceptive stimulus inhibition of, **264**
recruitment in, **261**

Extensor reflex, spinal cord section in man affecting, 234
Extensor thrust reflex, nature of, 231
Exteroceptor, definition of, 216
Extinction, conditioned reflexes showing, 326
  motor cortex showing, 307
Extracellular fluid, water in, 810
Extra-pyramidal paths, lesions of, 315
  motor cortex and, 309, 311
Extra-systole, nature of, **547**
Eye, accommodation in, 389
  adaptation in, 429
  chromatic aberration of, **393**
  colour responses of, 445
  focal length of, 387
  labyrinth reflexes affecting, 279
  optical system of, 383
  parasympathetic nerve supply to, 350
  peripheral aberrations of, 395
  pupil changes with convergence of, 378
  refraction of, 392
  refractive indices of media of, **387**
  resolving power of, **419**
  righting reflexes dependent on, **282**
  sections through, **384**
  sensation curve in, **425**, 426
  spherical aberration of, 394
  sympathetic nerve supply to, 347
  visual field of, 409, **410**
  X-ray measurements of, **399**
Eye movements, cortical control of, **299**, 300, 306
  control of, 374
  superior colliculi concerned in, 289
Eye muscles, innervation of, **375**
Eye reflexes, nervous pathways of, 289
Eyeball, blood supply to, **381**
  length, posterior nodal distance relation to, **401**
  muscles of, 373
  nerve supply to, **378**
Eyelids, structure of, 379
Eye-strain, hypermetropia causing, 397

Facilitation, cerebral cortex stimulation effecting, 321
  extra-pyramidal paths concerned in, **313**
  internuncial action potentials and, 247
  motoneurone showing, **243, 244**
  motor cortex showing, 307
Fæces, calcium in, 814
  composition of, 899
Fainting, cause of, 656
Fallopian tube, ciliary movement in, 169
  fertilisation in 1128
  ovum passage in, 1107
False vocal cords, function of, 465
Faraday-Tyndall phenomenon, nature of, 72
Faradic stimulation, effects of, 111
Fat, analysis of, 27
  bile secretion affected by, 884
  calorific value of, **791**
  carbohydrate formation from, 944
    in diabetes from, 967
  composition of, 25–26
  daily requirements of, 835
  emulsification of, 887
  essential in diet, 810
  gastric digestion of, 862
  glucose producing, 955
  growth needs of, **789**
Fat, heart usage of, 607
  intestinal absorption of, 909
  metabolism of, 939
    phospholipides function in, 950
    pituitary action on, 1087
    respiratory quotient showing amount of, **805**
    starvation effect on, 797
  oxidation of, 944
  pancreatic digestion of, 881
  placenta containing, 1135
  properties of, 26–27
  specific dynamic action of, 803
  starvation effect on usage of, **799**
  stomach emptying time affected by, 873
  thermal conductance of, 1039
Fat absorption, changes during, 951
Fatigue, skeletal muscle showing, **135**–136
  spinal reflexes showing, 262
Fatty acid, hydrotropy and intestinal absorption of, 909
  iodine values in different tissues of, **950**
  oxidation of, 944
Fatty degeneration, nature of, 949
Fechner's law, nature of, **370**, 371
Feet, temperature effect on temperature of, **1035**
  sweating on, 1031
Female, androgens effect on, 1126
  reproductive organs of, 1105
Femoral pulse, brachial comparison with, 635
Fenn effect, nature of, 126
Ferret, ovulation in, 1106
Ferri-porphyrins, nature of, 516
Ferritin, nature of, 520
Ferro-porphyrins, nature of, 516
Fertilisation, events in, 1099
Fertility, menstrual cycle and time of, 1130
Fever, basal metabolic rate affected by, 795
  course of events in, 1052
  thyroid gland activity in, 1072
  uric acid excretion affected by, 938
Fibrillation, nature of, 554
Fibrin network, structure of, **493**
Fibrin, properties of, 506
Fibrinogen, characters of, 529
  hæmorrhage effect on, 532
  properties of, 46, 506
Fibroin, nature of, 51
Fibrous protein, nature of, **43**
Fick principle, cardiac output measurement by, 568
Fillet, composition of, **285**
Filtration fraction, nature of, 1003
Final common path, nature of, **208**
Finger plethysmograph, pulse volume record by, **636**
Fish, accommodation in eye of, 388
  circulation arrangement in, 537
  lens in eye of, 388
Fixation reflex, nature of, 374, 376
Flavine adenine dinucleotide, action of, **99**
  co-enzyme action of, **94**
Flavine enzyme system, E_h of, 101
Flavo-protein, lactate to pyruvate reaction involving, 748
Fleisch differential stromuhr, principle of, **625,** 626
Flexor muscles, chronaxie of, 194
  cortical control of, 307
Flexor reflex, characteristics of, 250
  extensor reflex compared with, **261**–262

Flexor reflex, nature of, **230**
reflex arc concerned in, **213**
reticular formation stimulation effect on, **312**
spinal cord section in man affecting, 234
Flicker, cause of, 427
cortical potentials relation to, **321,** 322
light intensity affecting, 421
Flour, vitamin B_1 content of, 826
Fluoride, action on muscle of, 149
blood clotting affected by, 534
Focal length, X-ray measurement in eye of, **400**
Fœtal blood, oxygen content of, 1137
oxygen dissociation curve of, **1138**
Fœtal circulation, maternal circulation relation to, **1139**
Fœtus, blood cell formation in, 517
blood specific gravity in, 501
blood supply with growth of, 1135
circulation in, **1136**
glycogen storage in, 957
iodine value of fat in, 940
oxygen supply to, 1137
respiration in, 1139
Folic acid, extrinsic factor and, 520
nature of, 830
Food, alcohol absorption affected by, **840**
cardiac output affected by taking of, 573
gall bladder affected by, 885
urea and uric acid excretion following ingestion of, **939**
Food rationing, introduction of, 836
Foodstuffs, absorption of, 903
body storage of, **796**
calorific value of, **791**
composition of, 784
isodynamic replacement of, 806
percentage absorption of, 833
protein metabolism affected by, 800
vitamin A content of various, 822
vitamin content of, **829**
Foramen ovale, fœtal circulation through, 1137
Fovea, light intensity effect on vision at, **442**
spectral coefficient curves for, **444**
structure of, 404, **405**
structure of image at, 395
Foveal luminosity curves, notches in, **443**
Fractionation, reflex response showing, 255
Fragility, red blood cells showing, 509
Free energy, nature of, 67
Free-martin, nature of, 1157
Freezing point depression, osmotic pressure measurement by, 59
Frequency diagram, use of, **7**
Frog, circulation arrangement in, 537
heart anatomy in, 542
kidney, vascular supply to, **1007**
skin, permeability of, 82
urine formation in, **1005**
vago-sympathetic nerve stimulation in, **590**
Fröhlich's syndrome, nature of, **1084**
Frontal lobe, functions of, 300, **301**
Fructose, blood containing, 952
glycogen formation from, 959, 960
properties of, 32–33
semen containing, 1123
F.S.H., nature of, 1113
testis affected by, 1126
Fugacity, nature of, 101
Fumaric acid, structure of, **31**
Furane, nature of, **34**
Fuscin, function in vision of, 407

Galactose, properties of, 33
Gall bladder, food effect on, 885
function of, 882
Galvanic stimulation, effects of, 110
Galvani's experiment, nature of, **117**
Gametokinetic hormone, nature of, 1121
Gastric acidity, regulation of, 874
Gastric acids, determination of, 861
Gastric contents, emptying time and volume of, **873**
sampling of, 859
Gastric juice, actions of, 861
composition of, 859
composition of different types of, 868
humoral secretion of, 866
pernicious anæmia affecting, 520
secretion of, 862
Gastric movements, digestion and, 870
Gastric mucosa, electric potential of, 863
Gastric secretion, normal course of, 867
Gastrin, nature of, 867
Gastro-colic reflex, nature of, 897
Gastro-ileal reflex, nature of, 895
Gelatin, amino-acids in, 44
amino-acids lacking in, 808
iso-electric *p*H of, 74
properties of, 70
Gels, properties of, 70
Genes, independent association of, 1155
physical basis of, 19
Gennari, line of, 297, 304
Germ cells, formation of, 1095
Germinal vesicle, nature of, 1106
Gigantism, nature of, **1083**
Glass electrode, use of, 65
Glaucoma, causes of, 383
hypermetropia causing, 397
Gliadin, amino-acids in, 44
Globe thermometer, use of, 1047
Globin, amino-acids in, 44
properties of, 513
Globoglycoid, nature of, 507
Globular protein, structure of, 43
Globulin, electrolyte nature of, 76
properties of, 46
Globulin X, nature of, 139
Globus pallidus, functions of, 313, **314**
Glomerular capillary pressure, magnitude of, 1000
Glomerular filtrate, factors affecting rate of, 999
Glomerular filtration rate, measurement of, 1002
Glomerular fluid, composition of, 996
Glomeruli, function of, 995
structure of, 990
Glottis, nature of, 465
Glucoascorbic acid, action of, 830
Glucosamine, nature of, 34
Glucose, absorption rate from intestine of, 907
alkali action on, 35–36
body content of, 953
brain usage of, 670, 746
energy in oxidation of, 67
heart usage of, 607
metabolic fate of, **962**
oxidation of, 144
properties of, 32
red blood cell usage of, 510
renal threshold of, 963
skin storage of, 1032
T*m*, value of, **1009**

Glucose, tolerance curve, diet effect on, 972
    diabetes mellitus effect on, **977**
    nature of, **953**
  usage, diabetes effect on, 967
Glucose-1-phosphate, formation from glycogen of, **97**
Glucosides, varieties of, 35
Glucuronic acid, excretion of, 988
  nature of, 35
Glutamic acid, metabolism of, 930
  structure of, 41
Glutaminase, action of, 932
Glutamine, amino-acid synthesis from, 923
Glycerol, structure of, **30**
Glycerophosphoric acid, nature of, 25
Glycine, buffer action of, 65
  creatine formation and, 933
  structure of, 40
Glycocholic acid, nature of, 883
Glycocyamine, creatine formation and, 933
Glycogen, enzymic synthesis of, 94
  metabolic fate of, **962**
  phosphorylation of, **97**
  placenta containing, 1135
  properties of, 38, 957
  Purkinje fibres containing, 553
  skeletal muscle containing, 139
    contraction and disappearance of, 142
  sparing action on protein in starvation, 798
  tissue storage of, 957
Glycogenolysis, nature of, 958
Glycolipides, nature of, 30
Glycolysis, definition of, 142
  mechanism of, 955
  normal occurrence in blood of, 150
  reactions concerned in, 144
Glyconeogenesis, adrenal cortical hormone effect on, **973**
  anterior pituitary hormone effect on, 971
  nature of, 958
Glycoproteins, structure of, 50
Glycosides, nature of, 35
Glycosuria, causes of, 963
  diabetes mellitus causing, 977
  nature of, 955
G : N ratio, diabetes and, 944, 967
  phlorrhizin glycosuria and, 964
Goitre, causes of, 1074
Golgi apparatus, fat absorption and, 912
  nature of, **13**
  spermatozoa maturation and, 1099
Golgi cells, cerebellum containing, **333, 334**
  distribution in spinal cord of, 222
  types of, **202**, 204
Golgi network, nature of, **203**, 204
Goll and Burdach, composition of columns of, 224–**225**
Gonadotropic hormones, discovery of, 1087
  ovary affected by, **1112**
  pregnancy effect on excretion of, 1132
Gonads, adrenal action on, 1093
  adrenal gland relation to, 1063
  anterior pituitary interaction with, 1092
  thyroid interaction with 1093
"Goose-skin," nature of, 1028
Göthlin's fragility test, nature of, 639
Graafian follicle, structure of, **1105**
Granit's micro-electrode apparatus, **438**
Granulocytes, characters of, 523
Gravity, circulation affected by, 641, 683
Grey rami communicantes, composition of, 345
Growth, factors controlling, 21
Growth, food materials retained during, **789**
  protein requirement in, 802
  vitamin A effect on, **819**
Growth hormone, action of, **1083**
Growth rate, milk protein relation to, **1150**
Guanase, action of, 936
Guanine, structure of, **48**

Habit, basis of, 325
Hæm, nature of, 513
  structure of, **516**
Hæm pigments, chemical relationships of, **515**
Hæmatin, nature of, 513
Hæmatocrit, use of, **496**
Hæmatoporphyrin, nature of, 513
  structure of, **515**
Hæmin, nature of, 513
Hæmochromogen, nature of, 513
Hæmoconcentration, treatment of, 811
Hæmocyanin, molecular weight of, 45, 72
  nature of, 23
Hæmocytoblast, nature of, 517
Hæmocytometer, use of, 496
Hæmofuscin, nature of, 682
Hæmoglobin, absorption spectrum of, 512
  buffering power of, 733
  carbon dioxide reaction with, 737
  composition of, 47
  cyanosis with amount of, 772
  determination of, 497
  erythrocyte size relation to, 509
  high altitude effect on, **777**
  iron content of, 497
  iso-electric *p*H of, 74
  molecular weight of, 72
  oxygen combination with, **728**
  properties of, 511
  rate of regeneration of, 503
Hæmoglobinometer, use of, 497
Hæmokonia, nature of, 506
Hæmolysin, formation of, 704
Hæmolysis, nature of, 16, 509
Hæmolytic jaundice, nature of, 682
Hæmophilia, nature of, 535
  sex-linked inheritance of, 1158
Hæmopoietic factor, nature of, 520
Hæmopyrroles, nature of, **515**
Hæmorrhage, blood clotting in arrest of, 531
  blood volume restitution following, 495
  circulatory changes caused by, 680
  plasma protein formation after, 918
  tissue fluid movement after, 692
Hæmorrhagic states, causes of, 535
Hæmosiderin, function of, 813
  nature of, 522
  spleen containing, 682
Hair, factors affecting growth of, 1032
  tactile sensibility relation to, 478
Hair follicle, nerve plexus on, 216, **217**
  structure of, 1028
Hair growth, thyroid gland effect on, **1071**
Haldane, blood gas apparatus of, 497, **498**
Haldane-Pembrey respiration apparatus, use of, **787**
Hales, experiment of, 609
Halometer, use of, 507
Hamburger interchange, nature of, 733
Hand, sweating on, **1030**
  temperature of, **1035**
Haploid number, nature of, 1099
Haptophore group, nature of, 702

“ **Harden-Young ester,**” structure of, **34**
Hartridge's microstimulation apparatus, **441**
    theory of visual acuity, 420
Head, body tone affected by position of, **277**
    pose, reflexes concerned in, **280**-**281**
    sympathetic nerve supply to, 347
    temperature of, 1035
Head blood supply, regulation of, 658
Hearing, central nervous pathways of, **286**
    cortical representation of, 303
    mechanism of, 454
    organ of, 450
    phenomena of, 461
    reaction time of, 319
    theories of, 456
    tympanum function in 452
    Weber's law and, 370
Heart, acid effect on, **63,** 64
    adrenaline action on, **590,** 1066
    blood flow to, 629
    β-hydroxybutyric acid utilisation by, 948
    changes in form of, 561
    compensation by, 580
    contraction of, 550
    efficiency of, **576**
    excitatory wave spread in, **566**
    fat content of, 943
        iodine values in, **950**
    glycogen formation in, 607
    H ion concentration effect on, **583**
    lactate utilisation by, 150
    law of the, 576
    muscle. *See* Cardiac muscle.
    nervous regulation of, 583
    oxygen usage of, 576
        adrenaline action on, 591
        relation to diastolic volume of, 579
    parasympathetic stimulation effect on, 350
    perfusion of, **581, 582**
    pressure measurements in, **557**
    propagation of contraction wave in, 544
    pyruvate usage by, 746
    referred pain from, 354
    reflexes in, 592
    respiration effect on filling of, 567
    starvation and loss in weight of, **797**
    structure of, 548
    sympathetic nerve supply to, 348
    systolic output of, 567
    temperature effect on, 580
    “ tone ” of, 579
    work done by, 574
        in lifetime of, 541
Heart beat, neurogenic theory of, 543
    sequence of events in, 556
Heart block, nature of, 554
Heart disease, exercise affected by, **676**
Heart rate, refractory period relation to, 546
Heart sounds, cause of, 563
    graphic recording of, 563, **564**
Heart-lung preparation, use of, **571**
Heart-lung-kidney preparation, use of, 997
Heart-oxygenator preparation, **606**
Heart volume, arterial pressure change effect on, **577**
    venous return change effect on, **578**
Heat, acclimatisation to, 1051
    balance sheet, **1042**
        and its partition, 1041
    metabolic rate affected by, 1037
Heat cramp, cause of, 1052
Heat exhaustion, nature of, 1052
Heat loss, regulation of, 1038
Heat production, heat loss balance with, 793
    measurement in muscle of, **156**
    nerve showing, 185
    regulation of, 1036
Heat radiation, measurement of, 1047
Heat-regulating centre, location of, 1053
Heat regulation, breakdown of, 1051
Heat rigor, nature of, 126
Heat spots, distribution of, **475**
Heat-stroke, cause of, 1045
    limiting climatic conditions for, **1049**
Heavy nitrogen, amino-acid metabolism studied by, 920, 923
Hecht's theory of vision, 418
Helmholtz, theory of accommodation, 389
    theory of hearing of, 456
Hemiplegia, muscle responses after, 315
Henderson-Hasselbalch equation, 504
Henle's loop, function of, 1006
Henry's law, blood gases and, 727
Heparin, action of, 532, 534
    anaphylactic shock and, 706
    chemical nature of, 50
Hepatectomy, blood sugar affected by, **958**
Hepatic artery, portal vein connections with, 669
Hepatic circulation, 668
Heredity, conditioned reflexes and, 328
    factors concerned in, 1152
Hering-Breuer reflex, nature of, 760
Hering's theory of depth perception, 448
Herpes zoster, nature of, 235, 649
Heterozygous, meaning of, 1155
Hexokinase, action of, 144
    anterior pituitary and adrenal hormones and insulin action on, 974
    glycogenolysis and action of, 957
    glycolysis and action of, 955
Hexose phosphates, varieties of, **34**
Hexoses, structure of, 31
Hibernation, respiratory quotient in, 944
High altitude, Cheyne-Stokes breathing at, 771
    oxygen pressure at, **775**
Hippuric acid, excretion of, 986
Hirudin, action of, 534
His, bundle of, **551**
Histamine, anaphylactic shock and, 706
    blood vessels affected by, 662
    bronchi affected by, 721
    calcium effect on action of, 721
    circulation time measurement by, 623
    gastric secretion and, 867, **868**
        stimulation by, 859
    histidine origin of, **97**
        relation to, 42, 931
    nasal mucosa affected by, **709**
    shock relation to, 663
    skeletal muscle contraction liberating, 149
    uterine muscle affected by, **164,** 165
Histidine, histamine formation from, **97**
    metabolism of, 931
    structure of, 41
Histiocytes, inflammation and action of, 698
    nature of, 525
Histogram, nature of, 7
Histones, properties of, 45
History of physiology, 8
Histotoxic anoxia, nature of, 771
Hofmeister series, nature of, 74
Homatropine, eye affected by, 391
Homeostasis, nature of, 4

Homogentisic acid, nature of, 929
Homozygous, meaning of, 1155
Hormones, properties of, 1057
Horopter, nature of, 445
Houssay animals, nature of, 970
H-substance, nature of, 662
Human nerve, electrical stimulation of, **193**
Humanised milk, nature of, 1151
Humoral theory of nervous action, 360
Hunger, hypoglycæmia causing, 956
  nature of, 900
Hunger diabetes, nature of, 800
Hunger pangs, cause of, 874
Hyaluronidase, discus proligerus affected by, 1128
  spermatozoa containing, 1120
Hydræmic plethora, events in, 678
  lymph formation in, **690**
Hydrohcloric acid, gastric juice containing, 860
Hydrogen, balance sheet of, **788**
Hydrogen ion concentration. *See also* *p*H.
  meaning of, 63
  measurement of, 65–66
Hydrolases, actions of, 95, 96
Hydrosols, properties of, 71–73
Hydrotropic action, nature of, 883
Hydrotropy, absorption in intestine affected by, 909
Hydroxy-acids, amino-acid synthesis from, 923
Hydroxyglutamic acid, structure of, 41
Hydroxyproline, structure of, 41
Hypercalcæmia, effects of, 1077
Hypercapnia, nature of, 755
Hypercarbia, nature of, 755
Hypermetropia, causes of, 396
Hyperpnœa, nature of, 717
Hyperpyrexia, effects of, 1052
Hypertension, nature of, 661
Hyperthyroidism, effects of, 1074
Hypertonic solution, cerebro-spinal fluid affected by, 695
Hypodynamic heart, conditions in, 579
Hypogastric nerves, action on bladder of, **1019,** 1024
  composition of, 348
Hypogastric reflex, nature of, **349**
Hypoglycæmia, effects of, 956
  pituitary dysfunction and, 1085
Hypophysectomy, carbohydrate metabolism affected by, 969
  effects of, 1080
  lactation affected by, 1144
Hypophysio-portal circulation, nature of, 1079
Hypothalamus, anti-diuretic hormone control by, 1014
  frontal lobe connections with, **301**
  functions of, 233, 357
  motor cortex connections with, 311, **312**
  nuclei and connections of, **356**
  pituitary connections with, 1080
  sleep centre in, 331
  temperature regulating centre in, 1053
Hypotonia, cerebellar ablation causing, 339
  cerebellar injury in man causing, **340,** 341
Hypoxanthine, skeletal muscle containing, 140
  structure of, **48**
Hysteresis, nature of, 70

I.C.S.H., nature of, 1113
  pregnancy and excretion of, 1132
Identical twins, nature of, 1152
Idiomuscular contraction, nature of, 136
Ileocæcal valve. *See* Ileocolic sphincter.
Ileocolic sphincter, function of, 893
Ileum, food absorption in, **919**
Imbecillitas phenylpyruvica, nature of, 929
Imbibition, nature of, 70
Iminazol ring, structure of, **49**
Immunity, development of, 701
Impedance, nature of, 110
Inanition, basal metabolic rate affected by, 795
  metabolic rate affected by protein intake, 802
Incus, malleus relation to, **451**
Indican, excretion of, 987
  formation of, 930
Indicator yellow, nature of, 406
Indicators, *p*H measurement by, 65–**66**
Indol, formation of, 930
Inertia fly-wheel, use of, **134**
Infant, food intake of adult compared with, 1150
Infection, leucocyte action in, 699
Inflammation, course of events in, 697
  leucocytes function in, 525
Infra-red rays, eye insensitive to, 422
Inheritance, agglutinogens and, 502
  laws of, 1154
  morphology of, 1153
Inhibition, conditioned reflexes showing, 326
  cortical area concerned in, 328
  extra-pyramidal paths concerned in, **313**
  internuncial action potentials and, 247
  motonurones showing, 245
  recruitment and after-discharge in, **262**
  reflex response showing, 256, **257**
  theories of mechanism of, **260**
Injury, macrophage activity after, 697
Inosine, nature of, 936
Inositol, nature of, 140
Insensible perspiration, nature of, 1029
  water loss in, 811
Inspiration, muscles used in, 711
Inspiratory centre, location of, 750
Instinct, nature of, 323
Insulin, actions of, 968
  anterior pituitary extract effect on reactions to, **970**
  blood sugar effect on output of, 975
  mechanism of action of, 974
  molecular weight of, 45
Intercostal muscles, action in respiration of, 712
Intermittent claudication, nature of, 678
Internal capsule, basal ganglia relation to, **309**
  composition of, 309
  lesions in, 315
Internal inhibition, nature of, 327
Internuncial cells, action potentials in spinal cord of, **246–247**
Internuncial paths, types of, **242**
Interoceptor, definition of, 216
Intestinal blood vessels, adrenaline action on, 1065
Intestinal circulation, foodstuffs absorption and, **904**
Intestinal gas, formation of, 900
Intestinal juice, properties of, 888
Intestinal mucosa, enzymes in, 889
  phospholipides in, 950
Intestine, blood flow to, 629
  parasympathetic stimulation effect on, 350
  sympathetic stimulation effect on, 348
  See also Large and Small intestine.

Intra-abdominal pressure, respiration effect on, 715
Intra-auricular pressure, cardiac cycle and changes in, **561**
Intracellular fluid, water in, 810
Intracortical neurones, types of, **296–298**
Intracranial pressure, factors affecting, 671
Intra-ocular pressure, constancy of, 383
Intra-pleural pressure, cause of, 714
Intrathoracic pressure, cause of, 714
Intra-uterine pressure, parturition effect on, **1142**
Intravascular clotting, production of, 533
Intraventricular pressure changes,
  measurement of, 558
Intrinsic factor, anæmia and, 862
Inulin clearance, value of, 1003
Inulin, nature of, 38
Invertase, action of, 36
  fructose effect on, 92
Invertebrate hearts, structure of, 548
Involuntary muscle. *See* Plain muscle.
Iodine, daily requirement of, **812**
  thyroid gland uptake of, 1072
Iodine lack, thyroid gland affected by, 1073
Iodine value, body fat variation in, 939
Iodised casein, thyroxine formation from, 1072
Iodoacetate, action on muscle of, 148, 149
  fat absorption affected by, 913
  heart affected by, **607**
Ionophoresis, nature of, 1031
Iris, functions of, 377
  light sensitivity in some species, 377–378
Iron, absorption of, 895
  balance sheet of, **812**
  functions of, 23
  hæmoglobin containing, 497, 511
Iron wire model, nerve impulse and, 195
Irradiation, conditioned reflexes showing, 325
  flexor reflex showing, 251–252
  reflex action showing, 214
Ischæmic pain, nature of, 678
Islets of Længerhans, diabetes mellitus and, 978
  nervous control of, 1094
  structure of, **966**
Isochronism, law of, 115
Iso-citric acid, structure of, **31**
Isocortex, nature of, 293
Isodynamic replacement, foodstuffs and, 806
Iso-electric *p*H, proteins showing, 74
Iso-electric point, nature of, 40
Isohæmopyrrole, nature of, **515**
Isoleucine, structure of, 40
Isomerases, actions of, 96, 101
Isometric contraction, fibril changes with, **107**
  heat production in, 158
  nature of, **123–124**
Isosmotic solutions, definition of, 59
Isotonic contraction, fibril changes with, **107**
  heat production in, 158
  nature of, 121–**122**
Isotonic solutions, definition of, 59
sotopes, metabolic studies with, 920
Jaundice, blood clotting in, 535
  red blood cell fragility affected by, 509
  types of, 887
Jejunem, food absorption in, **919**
Joint movement, respiration affected by, 769
Joint, structure of, 818
Jugular pulse, heart block effect on, **554**
  heart sounds relation to, **564**
  radial pulse and electrocardiogram with, **565**
Jugular pulse, recording of, **640**
Jugular vein, pressure in, 639
Juxtamedullary glomeruli, nature of, 992

Kaes, line of, 303
Kallikrein, shock due to, 664
Katabolism, definition of, 3
Kephalin, structure of, 30
Kephir, nature of, 1148
Keratin, amino-acids in, 44
  properties of, 51
  structure of, **43**
Keto-acids, amino-acid metabolism and, 927
  amino-acid synthesis from, 923
  transamination and, 923
Ketogenic amino-acids, nature of, 928
Ketogenic hormone, action of, 1087
Ketoglutaric acid, structure of, **31**
Ketone bodies, metabolism of, 947
Ketosis, cause of, 946, 965
Kidney, adaptation to reduction of tissue of, 1011
  adrenal cortical deficiency effect on, 1060
  ammonia formation by, 932
  blood flow to, 629
  efficiency in diuresis of, 1012
  endocrine organs effect on, 1010
  fat iodine values in, **950**
  pituitrin action on, 1088
  structure of, **990**
  temperature effect on, **997**
  work done by, 994
Kinæsthetic impulses, nervous pathways for, **225**
Kinæsthetic sensation, central nervous
    pathways for, 284
  unconscious nervous pathways of, 292
Knee-jerk, cerebellar injury effect on, **340,** 341
  facilitation of, 265
  nature of, 230
  reflex arc concerned in, **213**
  reticular formation stimulation effect on, **312**
  vagus stimulation effect on, **359**
Knipping recording spirometer, use of, **786**
Korotkov sounds, nature of, 613
Krause's end bulb, function of, **481**
  nature of, 216, **217**
Krause's membrane, nature of, 106
Kupffer cells, nature of, 526
Kynurenic acid, nature of, 930

Labour pains, nature of, **1142**
Labyrinth, functions of, 217, 270
  structure of, **267**
Labyrinth reflexes, nature of, **278**
Labyrinthectomy, effects of, 270
Labyrinthine righting reflexes, head pose and, **280**
Lachrymal gland, nerve supply to, 350
  secretion of, **380**
Lactalbumin, milk containing, 1149
  molecular weight of, 45
Lactase, action of, 37
  galactose effect on, 92
  kinetics of action of, **91**–92
Lactation, control of, 1144
  reflex factors in, 1145
Lacteals, nature of, 686
Lactic acid, excretion of, 988
  exercise and production of, 150–**151**

Lactic acid, fatigue in muscle due to, 136
  gastric contents containing, 860
  heart usage of, 607
  liver removal of, 959
  pyruvate formation from, 748
  reduction rate of arterial blood affected by, **743**
  skeletal muscle containing, 139
    contraction and production of, 141
    heat production and changes in, 159
  structure of, **30**
  sweat containing, 1029
  vasomotor centre affected by, 652
Lactic acid cycle, isolated muscle and, 147
  nature of, 142
Lactic dehydrogenase, action of, **98**
  biliverdin affected by, 522
Lactoflavin, actions of, 827
Lactogenic hormone, action of, 1087
Lactoglobulin, milk containing, 1149
  molecular weight of, 45
Lactone ring, carbohydrates and, 33
Lactose, mammary gland synthesis of, 1147
  properties of, 37
  whey containing, 1149
Lævocardiogram, nature of, 566
Lanoline, nature of, 1028
Large intestine, functions of, 894
  innervation of, **896**
  movements of, 896
Laryngoscope, use of, 467
Larynx, anatomical mechanism of, 465
  deglutition and movement of, **855**
Latent addition, nerve showing, 180
  of excitation, plain muscle showing, 166
Latent period, plain muscle showing, 166
  reflex response showing, **252, 253**
  skeletal muscle showing, **122–123**
Lateral geniculate body, retinal area representation in, **288**
Lateral lemniscus, composition of, **286**
Law of forward direction, nature of, 205, **207**
  the heart, 576
  isochronism, nature of, 115
  plurality of connections, neurones and, **242**
  reciprocity of connections, neurones and, **242**
  specific irritability, 369
Lead poisoning, parathyroid effect on, 1078
Lecithin, fat absorption affected by, 912
  function in tissues of, 949
  structure of, **29**
Leech muscle, acetylcholine action on, **363**
Leg circulation, posture effect on, 641
Legumin, properties of, 46
Lengthening reaction, decerebrate preparation showing, 265
Lens, diffraction pattern formed by, 392
  focal length of, 387
  metabolism of, 385
  radius of, 386
  refractive power of, **385**
  structure of, 383
  suspension of, **384**
Lens capsule, thickness of, **390**
Lenticular nucleus, cortical connections with, 313, **314**
Leucine, structure of, 40
Leucocytes, amœboid movement of, 169
  chemiotaxis and, 3
  formation of, 524
  functions of, 525
  inflammation and action of, **698**
Leucocytes, number of, 493
  types of, 523
Leucotomy, results of, 301
Leukotaxin, nature of, 3, 699
L.H., nature of, 1113
Light, gonads stimulation by, 1113
  intensity threshold for, 413
  physical properties of, 372
  visual acuity affected by intensity of, 416
Light reflex, nature of, **290,** 377
Lillie's model, nerve impulse and, 195
Limb volume, adrenaline action on, **1065**
Limulus heart, structure of, 548
Linkage, heredity and, 1157
Linking, conditioned reflexes showing, 326
Linoleic acid, essential nature of, 949
  essential in diet, 810
Lipase, action of, 881
  bile salts action on, 887
Lipides, classification of, 25
Lipocaic, action of, 949
Lipoid, plasma membrane containing, **86**
Lipomicrons, nature of, 943
Lipotropic action, choline showing, 949
Liquor folliculi, nature of, 1106
Liver, adrenaline action on, 1066
  adrenaline inactivation by, 1068
  alcohol metabolism and, **841**
  amino-acid synthesis in, 923
  amino-acid uptake by, 916
  arginase, hormonal action on, 1086
  bile secretion in, 883
  blood flow to, 629
  capillary permeability in, 687
  choline action on, 949
  circulation in, 668
  deamination in, 926
  extirpation of, 669
  fat, glycogen relation to, 965
    iodine values of, **950**
    metabolism of, 947
  fatty degeneration of, 949
  fibrinogen formation in, 506
  function, test of, 882
  glycogen, adrenal cortical hormone effect on, **973**
    adrenaline action on, **961**
    bilirubin formation affected by, 522
    factors regulating, 976
    lactate as source of, 151
    regulation of, 956
  hæmopoietic factor of, 520
  insulin action on, 974
  jaundice in absence of, 887
  nervous control of, 963
  nucleic acid metabolism in, 937
  prothrombin formation in, 530
  serum proteins formation in, 506
  size, factors affecting, 881
  specific dynamic action of protein in, 803
  sugar output by, **959**
  sympathetic stimulation effect on, 348
  thyroxine formation by, 1073
  urea formation in, 926
  vitamin A content of, 821
  vitamin K action on, 825
Local excitatory state, nature of, 180
Local sign, nature of, 479
Locomotor ataxy, nature of, 316
Long sight. *See* Hypermetropia.
Loudness, scale of, 464

Loudness spectrum, outer and inner ear comparison, **461**
Lower motor neurone, acetylcholine synthesis in, 364
  effects of lesions of, 315
Lumbar cord, defæcation centre in, 899
Lumbar puncture, nature of, 693
Luminosity curve, light intensity effect on, **413**
Lung capillaries, distensibility of, 665
Lung inflation, respiration affected by, 759
Lungs, blood glycolysis in, 605
  dust removal in, 779
  gas exchange in, 738
  glycolysis in, 955
  glycolytic action of, 150
  oxygen diffusion in, 740
  oxygen uptake by blood in, 730
  respiratory movements of, 712
  speech control by, 471
  structure of, **708**
  sympathetic nerve supply to, 348
  volume of, 716
  water loss from, 811
Luteal cells, nature of, 1107
Lutein, nature of, 1107
Lymphagogues, action of, 689
Lymphatic system, blood relation to, **685**
  fat absorption by, 910
Lymphatic vessels, valves in, **691**
Lymph, composition of, **686**
  formation of, 685
  history of, 488
  movement of, 690
  plasma protein content comparison with, **687**
Lymph glands, functions of, 691
  lymphocyte formation in, 525
Lymphocytes, characters of, 524
  infection and action of, 700
Lysine, structure of, 41
Lysins, formation of, 704
Lysozyme, nature of, 697

Macrophage, action of, 697
Macula lutea, cortical representation of, 305
Maculæ acousticæ, nature of, 268, 402
Magnesium, acetylcholine synthesis and, 364
  balance sheet of, **812**
  excretion of, 985
  function in diet of, 813
  phosphatase affected by, 818
Magnesium sulphate, blood clotting affected by, 534
Male, œstrogens effect on, 1126
  reproductive organs of, 1119
Malic acid, structure of, **31**
Malleus, incus relation to, **451**
Malphighian body, nature of, 990
Malphighian follicles, nature of, 681
Maltase, action of, 36, 89
  pancreatic juice containing, 881
Maltose, properties of, 36
Mammary gland, corpus luteum effect on, 1113
  hormonal control of, 1104
  œstrogens effect on, 1111
  pituitary action on, 1087
  pregnancy effect on, 1144
Mammogen, nature of, 1144
Man, composition of, **789**
Manganese, function in diet of, 813
Mannose, properties of, 33
Marey's law, nature of, 595
" **Mark-time** " reflex, mechanism of, 260–**261**
Mass peristalsis, nature of, 897
Mass reflex, nature of, 234
Mast cells, characters of, **524**
  heparin formation in, 533
Mastication, process of, 852
  reflex nature of, 261
Mayer waves, nature of, 653
Meat, composition of, 835
  gastric secretion response to, 869
Mechanical efficiency, exercise and, 804
Mechanism, meaning of, 6
Medulla, electric potentials in, 321
  respiratory centres in, 749
  structure of, 210, **211**
Medullary sheath, structure of, 172
Megakaryocyte, nature of, 527
Megaloblast, nature of, 519
Meiosis, cancer cells showing, 1153
  nature of, **1097**
Meissner's corpuscle, function of, **481**
  nature of, 216, **217**
Melanin, skin containing, 1028
Melanophores, pituitary action on, 1088
Membrana granulosa, nature of, 1106
Membrane manometer, endocardiac pressure measurement by, **557**
Membrane equilibria, factors concerned in, 82–84
Membrane, permeability of, 80–82
Membrane potentials, nature of, 85
Mendel's laws, 1154
Menopause, changes at, 1119
  gonadotropic hormone excretion following, 1133
Menstrual cycle, events during, **1116**
  fertile period in, 1130
Menstruation, causation of, 1116, 1118
  nature of, 1115
  olfactory sense affected by, 487
Mercurial manometer, use of, **610**
Mercury, skin absorption of, 1031
Mesial lemniscus, composition of, 284, **285**
Metabolic rate, exercise effect on, **804**
  protein intake effect on, 802
  starvation effect on, **797**
  vitamin requirements relation to, 820
Metabolism, definition of, 3
  history of, 782
  specific dynamic action in, 802
Metachronal wave, nature of, 170
Metallo-porphyrins, nature of, 516
Metals, colloidal solutions of, 69
Metaproteins, nature of, 46–47
Methæmalbumin, nature of, 523
  properties of, 512
Methionine, creatine formation and, 933
  liver affected by, 808
  lipotropic action of, 949
  metabolism of, 931
  structure of, 40
Methylene blue, tissue respiration and, 747
Methylglyoxal, structure of, **30**
Meynert, giant pyramidal cells of, 304
Microcalorimeter, muscle heat production measurement by, **157**
Micro-manipulator, use of, **12**
Micro-pipette, glomerular fluid sampling by, 996
  tubular fluid sampling by, **1004**
Microrespirometer, use of, **145–146**
Microrespirometry, use of, 745

Microtonometer, blood gas pressures measurement by, 739
Micturition, physiology of, 1016
  spinal cord section affecting, 234
Mid-brain animal, brain section producing, **273**
Mid-brain, micturition centre in, 1025
Milk, clotting of, 1149
  composition of, 1148
  cow and human comparison, **1151**
  fat, composition of, 26
   origin of, 940, 1147
  gastric juice action on, 862
  pancreatic juice action on, 880
  protein, young growth rate relation to, **1150**
  secretion of, **1147**
  vitamin A content of, 822
Miner's cramp, cause of, 812
Mitochondria, nature of, **12,** 13
  spermatozoa maturation and, 1099
Mitotic division, process of, **1095**
Mitral valve, structure of, 550
Modiolus, nature of, 453
" **Modulator,"** nature in vision of, **440**
Molecular weight, glomerular membrane permeability and, 998
Monocular depth perception, 446
Monocytes, characters of, 524
Mono-molecular film, nature of, 79
Monophasic action potential, cause of, 119
Morgagni, ventricle of, 465
Morphine, pupil affected by, 379
  vagal tone affected by, 588
Morula, nature of, 1130
Moss fibres, origin of, 335
Motoneurone pool, nature of, 250
Motoneurones, constant current stimulation of, 247, **248**
  facilitation in, **243, 244**
  transmission at, **240–241**
  volley rate of, 246
Motor areas, ablation of, 308
  location of, 305
  localisation in, **299,** 300
  stimulation of, 306
Motor cortex, extra-pyramidal paths from, 309
  hypothalamic stimulation effect on response to stimulation of, **312**
  suppressor strip effect on, **314**
Motor end plate, transmission at, **196**
  structure of, **106,** 108
Motor nerve, composition of, 108
  terminations of, 108
Motor paths, lesions in man of, 315
Motor point, stimulation of, 193
Motor unit, functioning of, 130
  nature of, 204
  size of, 172–173, 250
   in different muscles of, 306
Mountain sickness, nature of, 774
Mouth, digestion of, 842
  speech affected by muscles of, 471
  temperature in, 1035
Movement, nervous pathways for conscious sensation of, 224–**225**
Mucin, gastric digestion of, 862
  properties of, 50
Mucinogen, salivary gland containing, 851
Mucoid, properties of, 51
Mucoitin, nature of, 50
Mucus, gastric secretion of, 868
Müller, retina containing fibres of, 404
Müller's muscle, nerve supply to, 348
Multiple chain, neurone action and, 244–245
Multiple cone hypothesis of colour vision, 437
Murmur, physical cause of, 617
Muscarine, action of, 351
Muscular movements, afferent impulses affecting co-ordination of, 316
  cortical stimulation giving, 306
Muscular tone, semi-circular canals concerned in, 272
  utricle concerned in, 271
Musculus vocalis, action of, 467
Music, vibration frequencies used in, 463
Mutarotation, nature of, 33
Mutases, action of, 101
Mutation, heredity and, 1159
Myasthenia gravis, nature of, 197
  thymus in, 1091
Myelin sheath, structure and function of, 172, 173
Myenteric reflex, nature of, **891**
Mylohyoid muscle, preparation of, 105
Myofibril, structure of, **107**
Myogen, properties of, 46, 138–139
Myoglobin, function of, 107
  oxygen dissociation curve of, **744**
Myopia, causes of, 398
Myosin, adenosine triphosphatase and, 148
  muscle contraction effected by, **160**–161
  physical properties of, 107–108
  properties of, 46, 139
Myxœdema, nature of, 1070

Narcotics, nerve affected by, 186
Nasal fossæ, olfactory epithelium in, **486**
Nasal passages, factors affecting diameter of, **709**
Neck, sympathetic nerve supply to, 347
Neck reflexes, eye movements affected by, 280
  righting reflexes and, **281**
  types of, **277**
Necturus, urine formation in, **1005**
Negative successive induction, nature of, 254
  spinal reflexes showing, 260
Neocerebellum, composition of, **332, 333**
  effects of injury in man to, **341**
  effects of removal and stimulation of, 339
Nerve, accommodation time factor in, 192
  action potentials in, **120, 175**
  block, excitability changes in nerve below, **187**
  conduction, electrotonic changes in, 189
  conductivity, factors affecting, 186
  degeneration of, 173
  electrotonic potentials in, 190–**191**
  electrotonus in, **189–190**
  excitability, electrotonic changes in, 189–**190**
   parathyroidectomy effect on, 1076
   temperature effect on, 181
  excitation of, 174
  fibres, classification of, 183
  history of, 103
  impulse, acetylcholine concerned in, 365
   all-or-none character of, 180
   chemical changes accompanying, 185
   frequency of, 178
   frequency, reflex contraction and, 246
   heat production in passage of, 185
   nature of, **194**–195
   train, nature of, 178
   transmission speed of, 177, **182**

Nerve, impulse, velocity in different fibres of, 177, **183**
volley, nature of, 178
interaction between adjacent fibres of, **184**
refractory period of, **179**
regeneration of, 173
structure of, **171**
terminations of, **202,** 204
theory of impulse transmission in, 185
unfatiguability of, 187–**188**
velocity of impulse in different fibres of, 177, **183**
Nerve centre, definition of, 212
Nerve tracks, spinal cord section and degeneration of, **222,** 223
Nerve-muscle preparation, chronaxie of, 114
use of, 105
Nervi erigentes, functions of, 350
large intestine secretion control by, 895
Nervous system, chemical transmission in, 360
“ **Neuberg ester,**” structure of, **34**
Neurilemma, nature of, 172
nerve regeneration and, 173
Neurocyton, structure of, **203**
Neurofibrils, nerve cells containing, **203,** 204
Neurokeratin, properties of, 51
Neurology, conditioned reflexes relation to, 329
Neuro-muscular conduction, thymus effect on, 1091
Neuro-muscular spindle, response to stimulation, 219–**220**
structure of, **217–218**
Neurone pool, nature of, **211,** 212
Neurone, types of, **202–203**
Neurosis, experimental development of, 327
New-born, composition of, **789**
renal function in, 1011
Niacin, actions of, 827
Nicotinamide, actions of, 827
Nicotine, sympathetic ganglia affected by, 345, 351, 645
sympathetic ganglia potentials affected by, 240
Nicotinic acid, actions of, 828
Nictitating membrane, nature of, 381
sympathetic nerve supply to, 347
sympathetic ganglion stimulation effect on, 238
Night-blindness, causes of, 434
vitamin A concerned in, 407
vitamin A, effect on, 821
Nissl bodies, nature of, **203**
Nitric oxide, hæmoglobin reaction with, 511
Nitrogen, absorption coefficient of, **726**
balance sheet of, **788**
excretion, starvation effect on, **799**
fixation of, 22–23
solubility in tissues of, 778
Nitrogenous equilibrium, factors affecting, 800
intravenous amino-acids injection and, 915
nature of, **785**
Nociceptive stimulus, prepotent influence of, **258**
Node of Ranvier, nature of, 172
Nomogram, effective temperature from, **1048**
serum calcium and protein, **814**
Non-polarisable electrodes, use of, 116, **117**
Non-protein respiratory quotient, carbohydrate and fat utilisation shown by, **805**
fat and carbohydrate usage measurement by, **798**
Nor-adrenaline, actions of, 1065
sympathin comparison with, 361
vasoconstriction action of, 659
Norleucine, structure of, 40
“ **Normal,**” definition of, 6
“ **Normal saline,**” definition of, 61
Normoblast, nature of, 519
Norvaline, structure of, 40
Novocaine, nerve affected by, 186
Nuclear delay, nature of, 243
Nuclease, action of, 935
Nucleic acid, breakdown of, 935, **937**
composition of, 48
molecular weight of, 72
Nucleins, composition of, 47
formation in body of, 935
properties of, 48
Nucleoproteins, composition of, 47
gastric digestion of, 862
metabolism of, 935
molecular weight of, 50
properties of, 48
Nucleosidase, action of, 936
Nucleoside, structure of, 49
Nucleotidase, action of, 935
Nucleotide, structure of, 49
Nucleus, cell division affecting, 1096
composition of, 17, **18**
effect of damage to, **18, 19**
structure of, **9,** 12
Nucleus cuneatus, function of, 284
Nucleus gracilis, function of, 284
Nutrition, history of, 781
Nystagmus, cerebellar lesions causing, 341
rotatory movement of head producing, 283
semi-circular canal stimulation inducing, 27

Obesity, chylomicrograph affected by, **910**
Occipital cortex. *See* Visual cortex.
Occipital lobe, functions of, 301
Occlusion, reflex response showing, **256**
Occlusion test, nature of, 662
Oculo-motor nerve, effect of injury to, 375
Oculo-motor neurones, transmission at, **240–241**
Oculo-motor nuclei, connections of, **375**
Œdema, cause of, 689
Œsophagus, deglutition and action of, **857**
parasympathetic stimulation effect on, 350
Œstradiol, structure of, **1110**
Œstriol, structure of, **1110**
Œstrogens, anterior pituitary affected by, 109
capon’s comb affected by, **1124**
male excretion of, 1126
mammary gland affected by, 1144
menstrual cycle effect on excretion of, 111
nature of, 1110
parturition effected by, 1141
placenta formation of, 1134
pregnancy effect on excretion of, **1133**
prostate affected by, 1122
Œstrone, pregnancy test by injection of, 113
structure of, **28, 1110**
Œstrus, causation of, 1109
vaginal changes in, **1108**
Œstrus cycle, course of, **1108**
Ohm’s law, sound analysis and, 463
Oleic acid, body synthesis of, 949
depôt fat containing, 939
nature of, 26
Olfactometry, use of, 487

lfactory epithelium, structure of, 486
lfactory sensation, 485
live, cerebellar connections with, 335
livo-cerebellar tracts, connections of, 292
ögenesis, process of, 1099
‘ **Opening tetanus,**’’ nature of, 189
perative temperature, nature of, 1040
phthalmoscope, use of, **411**
psonins, nature of, 705
ptic disc, appearance of, 411, **412**
ptic nerve, composition of, 287–288
ptic paths, course of, **287**
ptic radiation, blindness resulting from damage to, 289
Optic tract, blindness resulting from damage to, 289
Orchidectomy, effects of, 1103
Organ of Corti, nature of, **455**
“ **Organiser,**” nature of, 21
Orgasm, nature of, 1127
Ornithine, structure of, 41
  urea formation and, **927**
Orthopnœa, nature of, 717
Oscillograph, use of, **118,** 119
Osmometer, colloid osmotic pressure measurement by, **71**
Osmoreceptors, nature of, 1014
Osmosis, absorption from intestine effected by, 906
Osmotic diuretics, action of, 1006
Osmotic pressure, measurement of, 58–**61**
  nature of, **57**
Ossein, nature of, 815
Osteoclast, nature of, 526
  parathyroid effect on, 1078
Otolith organ, nature of, 268, **271**
Ovariectomy, effects of, 1103
Ovary, history of, 1056
  menstrual cycle changes in, **1117**
  pregnancy effect on, **1131**
  structure of, **1106**
Over-ventilation, apnœa following, 767, **770**
Ovulation, process of, 1105
  time of, 1117
Ovum, fertilisation effect on activity of, 1101
  fertilisation of, 1128, 1130
  formation of, 1099
  maturation of, 1106
Oxalate, blood clotting affected by, 528
Oxalic acid, excretion of, 987
Oxalo-acetic acid, fat metabolism and, 947
  pyruvic acid oxidation and, **143**
  structure of, **31**
Oxalosuccinic acid, structure of, **31**
Oxidases, actions of, 96, 99
  high oxygen pressure effect on, 774
Oxidation-reduction potentials, nature of, 101
Oxidising enzymes, action of, 98
Oxoisomerase, action of, 101
Oxycalorimeter, use of, **790**
Oxygen, absorption coefficient of, **726**
  blood content and saturation with, 500
  calories per litre of, **791**
  carbon dioxide dissociation curve affected by, **734**
  coefficient of utilisation in exercise, 676
  diffusion constant of, 740
  Fick principle use of, 569
  red blood cell formation affected by pressure of, 520
  tissue usage of, 745
Oxygen consumption, exercise affecting, **152**
Oxygen consumption, exercise effect on, 673
  isolated muscle showing, **146**
  of heart, work effect on, **576**
  protein meal effect on, **916**
  temperature effect on, 1037
Oxygen debt, exercise and, 152–**153**
  isolated muscle showing, 146
Oxygen dissociation curve, carbon dioxide effect on, **728**
  carbon monoxide effect on, 779
Oxygen insufflation, efficacy of, 719
Oxygen lack. *See* Anoxia.
Oxygen pressure, air and blood comparison, **740**
  effects of high, 773
  high altitude effect on, **775**
Oxyhæmoglobin, crystals of, **511**
  factors affecting dissociation of, **743**
  formation of, 511
  iso-electric pH of, 74
Oxyntic cells, composition of secretion by, 868
  secretion by, 863
Oxytocin, action of, 1088
  lactation affected by, 1146
  parturition effected by, 1141
  uterus and cervix affected by, 1141

Pace-maker, action of, **553**
  nature of, 543
Pacinian corpuscle, function of, **481**
  nature of, 216, **217**
  vasodilator reflex from stimulation of, 656
Pain, central nervous pathways for, 284
  pupil size in, 379
  nerve tracts concerned in sense of, **225,** 226
  sense of, 480
  spino-thalamic tract section and, 235
Palæocerebellum, composition of, **332,** 333
  effects of injury in man to, **340**
  effects of removal of, 339
Palm, sweat secretion on, **1030**
Palmitic acid, nature of, 26
p-**aminobenzoic** acid, sulphanilamide competition with, 820
p-**aminohippuric** acid Tm, value of, 1009
Pancreas, blood sugar action on, 1091
  blood sugar effect on insulin secretion by, 975
  diabetogenic hormone action on, 971
  insulin content of, 969
  parasympathetic stimulation effect on, 350
  zymogen granules in, **878**
Pancreatectomy, effects of, 965
  hypophysectomy effect on results of, 970
Pancreatic diabetes, nature of, 965
Pancreatic fistula, nature of, 875
Pancreatic juice, diet effect on composition of, **881**
  properties of, 878
  secretion of, 875
Pancreatic secretion, histological changes with, **878**
Pancreozymin, nature of, 877
Pantothenic action, nature of, 827
Papillary muscles, function of, 550
Paracasein, nature of, 1149
Paraffin, intestinal absorption of, 910
Paraganglia, nature of, 1060
Parallax, monocular depth perception use of, 447
Paralytic ileus, nature of, 893
Paraplegia, types of, 234

Parasympathetic nerve stimulation, salivary gland response to, **847**, 851
Parasympathetic system, chemical transmission in, 361
  composition of, 350
  hypothalamic stimulation effect on, 357
  irradiation in, **359**
  sympathetic comparison with, 351
Parasympathetic vasodilator nerves, 646
Parathyroid glands, function of, 1076
  history of, 1055
  lactation affected by, 1145
Parathormone, effects of injection of, 1077
Paraxanthine, thyroid gland affected by, 1076
Parietal cortex, structure of, **297**
Parietal lobe, functions of, 300
Parkinson's disease, nervous pathways affected in, 313
Parthogenesis, nature of, 1101
Partitional calorimetry, nature of, 1041
Parturition, events in, 1140
  nerve centre for, 1142
  oxytocin function in, 1090
"Past-pointing test," nature of, 272
Patellar reflex, hypothalamic stimulation effect on, **312**
Pavlov pouch, use of, **864**
Pecten, accommodation in eye of, 388
Pecten muscle, " catch action " of, 165
Pectin, absorption of, 894
Pellagra, nature of, 827
Pelvic visceral nerves, action on bladder of, **1019**, 1024
Penis, coitus and action of, 1127
  function of, 1120
Pentose, nucleic acid containing, 49
  structure of, **31**
Pepsin, action of, 39, 861
Pepsinogen, histamine and vagal stimulation effect on, **869**
  properties of, 863
Peptic cells, secretion by, 863
Peptide, synthesis of, 42
Peptide linkage, nature of, 42
Peptone, blood clotting affected by, 534
  feeding, blood amino-acids after, **917**
  properties of, 47
Pericardial fluid, coagubility of, 534
  effect of excess of, 567
  function of, 550
  plasma composition comparison with, **692**
  serum composition comparison with, **687**
Pericardial pressure, respiration effect on, 668
Pericardium, function of, 550, 567
Perimeter, visual fields measurement by, **409**
Perineuronal spaces, nature of, 695
Peripheral resistance, measurement of, 617
Peripheral vision, extent of, 409
Peristaltic movement, mechanism of, **891**
Peritoneal cavity, fluid absorption from, 692
Peritoneal fluid, plasma composition comparison with, **692**
Peritoneum, referred pain from, 354
Perivascular spaces, function of, 694
Pernicious anæmia, gastric juice in, 860
  treatment of, 520
Peroxidases, action of, 100
Personality, cortical areas concerned in, 301
Petechiæ, nature of, 639
P-factor, ischæmic pain and, 678
"Pfluger's law of contraction," nature of, **190**
*p***H**, definition of, 63
*p***H**, nerve affected by, 186
  plain muscle affected by, **164**, 166
  red blood cell fragility affected by, 509
  respiratory centre affected by, 765
Phagocytosis, nature of, 525, 697
Pharynx, deglutition and movements of, **855**
Phase contrast, use of, 11
Phenol red, renal secretion of, 1007
Phenylalanine metabolism of, 928
  structure of, 40
Philippson's reflex, nature of, **265**
Phlorrhizin, action on kidney of, **1005**, 1010
Phlorrhizin glycosuria, mechanism of, 964
Phosphatase, bone formation and, **817**
  fat absorption in intestine affected by, 909, 912
Phosphate, bone formation and, 816
  buffer action of, 64
  daily requirement of, **812**
  energy-rich bond of, 52
  excretion of, 895, 985
  fat formation from carbohydrate and, 941
  glucose absorption from intestine affected by, 907
  metabolism, parathyroid effect on, 1078
    vitamin D action on, 823
Phosphoacetic acid, structure of, **53**
Phosphoglucomutase, glycogenolysis and action of, 957
Phosphoglyceric acid, structure of, **31**
Phospholipides, function in tissues of, 949
  nature of, 29
Phosphoproteins, properties of, 47
Phospho-pyridino-nucleotides, co-enzyme action of, 93
Phospho-pyruvic acid, structure of, **30**, **53**
Phosphorus, functions of, 23
Phosphorylases, actions of, 95, **97**
  glycogenolysis and action of, 957
Photo-chemical reaction, types of, 373
Photo-chromatic interval, nature of, 415
Phthalate, buffer action of, 65
Phyllopyrrole, nature of, **515**
Physiology, definition of, 1
  history of, 8
  method of, 5
Physostigmine, eye affected by, 391
Phytin, nature of, 814
  rickets production by, 823
Pia mater, innervation of vessels in, 671
  nature of, 693
Piezometric circulation, nature of, **618**
Pilocarpine, action of, 351
Pilomotor nerves, course of, 346
Pinnaglobulin, nature of, 24
Piqûre glycosuria, mechanism of, 964
Pituicytes, nature of, 1079
Pituitary basophilism, signs of, **1086**
Pituitary body, structure of, **1079**
Pituitary gland, history of, 1055
  nervous control of, 1093
Pitressin, coronary vessels affected by, 602
  nature of, 1088
Pituitrin, actions of, 1087
  sweating effected by, 1031
Placenta, functions of, 1134
  mammary glands affected by, 1144
  œstrogen formation by, 1111
  pregnancy maintenance by, 1141
  uterine blood content and growth of, **1134**
Placental circulation, fœtal circulation relatio to, **1139**

“Placing reaction,” nature of, 283
Plain muscle, action potentials in, 167–**168**
  chemical agents affecting, **164**, 166
  chronaxie of, 114
  contraction, chemistry of, 169
  pituitrin action on, 1088
  propagation of excitatory state in, 168
  properties of, 162
  stimulation of nerve to, **165**
  temperature effect on oxygen usage by, 1036
  tonus in, 163–165
Plantar reflex, spinal segments concerned in, 215
Plasma, carbon dioxide dissociation curve of, **732**
  carbon dioxide solubility in, **727**
  clearance, meaning of, 1002
  composition of, **505**
  definition of, 493
  globulin, agglutination and, 501
  osmotic pressure, red blood cell fragility affected by, 509
  oxygen solubility in, **727**
  potassium, storage of blood effect on, 510
  proteins of, 506
    glomerular filtration affected by concentration of, 1002
    protein metabolism and, 917
    renal excretion of, 998
    starvation effect on, 800
  urine composition compared with, **995**
  vitamin A content of, 821
  volume of, 496
  water of body in, 810
Plasmapheresis, nature of, 918
Plasmolysis, nature of, 15, **16**
  osmotic pressure measurement by, 59
Plastein, nature of, 94
Platelets, characters of, **526**
  number of, 493
Plethora, events in, 678
Plethysmograph, blood flow measurement by, 627, **644**
  pulse volume measurement by, **636**
Pleura, referred pain from, 354
  structure of, 708
Pleural cavity, gas absorption from, 741
  fluid absorption from, 692
Pleural fluid, coagubility of, 534
P.M.S., nature of, 1133
Pneumotaxic centre, action of, **753**
  nature of, 751
Pneumothorax, nature of, 714
Poiseuille's formula, blood flow and, 616
Polar body, nature of, 1099
Polarisation, nature of, 80
  nerve impulse and, 185
Polarising current, nature of, 189
Polar stimulation, nature of, **110**
Polocyte, nature of, 1099
Polychromatic theory of colour vision, 439
Polygraph, use of, **632**
Polypeptides, properties of, 47
  trypsin action on, 880
Polysaccharides, nature of, 37
Pons, structure of, **211**–212
Porphin, structure of, 515
Porphyrin derivatives, skeletal muscle containing, 140
Porphyrins, excretion of, 987
  structure of, **515**
Portal pressure, exercise effect on, 680
Portal pressure, plethora effect on, **679**
Portal reservoir, factors affecting, 669
Portal vein, pressure in, 668
Posterior root ganglion, autonomic cells in, **352**
Post-ganglionic fibre, composition of, 345
Post-pituitary gland, functions of, 1089
  hormones secreted by, 1087
  nervous connections of, **357**
  water excretion regulation by, 1013
Postural albuminuria, nature of, 998
Postural reflexes, nature of, 274
Postural tonus, cerebellar lesions effect on, 340
  nerve paths concerned in, **274**
Posture, alveolar air composition affected by, 724
  blood volume variation with, 495
  cardiac output affected by, 573
  circulation changes with, 684
  factors concerned in, 267
  leg circulation affected by, 641
  stretch reflex importance in, 264
  systems concerned in, **275**
  tissue fluid formation affected by, 688
Potassium, acetylcholine release affected by 366
  acetylcholine synthesis and, 364
  adrenal cortex affected by, 1062
  cardiac muscle affected by, **581**
  cell content of, 81–**82**
  daily requirement of, **812**
  excretion of, 985
  nerve affected by, 186
  nerve impulse and, 195
  skeletal muscle affected by, 138
    containing, 140
  starvation effect on excretion of, 799
P-P factor, actions of, 827
Precipitin, formation of, 704
Pre-ganglionic fibre, composition of, 345
Pregnancy, blood sedimentation rate affected by, 501
  blood oxygen dissociation curve affected by, **1138**
  course of, 1130
  hormone factors in, 1132
  hormonal events in, 1104
  ovariectomy effect on, 1115
  oxytocin action affected by, 1088
  pituitary changes during, 1079
  placenta functions in, 1134
  test for, 1132
  white cell count in, 524
Pregnancy urine, œstrogens in, 1110
Pregnandiol, placenta formation of, 1134
  pregnancy effect on excretion of, 1133
  structure of, **1114**
Presbyopia, nature of, 391
Pressor reflex, nature of, 654
Pressure diuresis, nature of, 1012
Presystolic bruit, cause of, 564
Primordial follicles, nature of, 1106
Probable error, definition of, 7
Progeria, nature of, **1081**
Progesterone, adrenal cortex containing, 1061
  mammary gland affected by, 1144
  menstrual cycle effect on excretion of, 1118
  pregnancy maintenance by, 1141
  structure of, **1114**
Progression, stato-kinetic reflexes concerned in, 283
Prolactin, action of, 1087
  lactation control by, 1145

Prolactin, mammary glands affected by, 1144
  œstrogen stimulation to formation of, 1134
Proline, metabolism of, 930
  structure of, 41
Pro-œstrus, vaginal changes in, **1108**
Proprioceptive impulses, cerebellar recording of, 335–**337**
Proprioceptive reflex, cerebellar injury effect on, 340
  characteristics of, 263–**264**
Proprioceptors, nature of, 216–**218**
Prostate, androgens effect on, **1125**
  secretion of, 1122
  structure of, 1119
Prostigmine, nature of, 197
Protein, amino-acids in, 43–44
  bile secretion affected by, 884
  biological value of, 807
  buffer action of, 65
  calorific value of, **791**
  catabolism of, 924
  chemical nature of, 38
  denaturation of, 77
  electrical charge on, 73
  electrolyte nature of, 76
  fat formation from, 942
  functions in diet of, 807
  growth needs of, **789**
  intestinal absorption of, 914
  lens containing, 385
  lipotropic action of, 949
  metabolism, dietary protein effect on, 800
    exercise effect on, 805
    growth hormone effect on, 1084
    plasma proteins action in, 918
    starvation effect on 797
  minimum requirements for, 834
  molecular weight of, 44–45
  pepsin action on, 861
  physiological heat value of, 791
  rate of intestinal absorption of, **916**
  renal excretion of, 998
  sparing action of glycogen in starvation, **798**
  specific dynamic action of, 802
  specificity of, 51
  storage in body of, **799, 801**
  synthesis in body of, 922
  ultra structure of, **43**
Proteoses, properties of, 47
Prothrombin, characters of, 529
  vitamin K effect on formation of, 825
Protamines, properties of, 45
Protopathic sensation, nature of, 481
Protoporphyrin, nature of, 516
Protoplasm, chemical constituents of, 22–24
  nature of, 10
  physical structure of, 14
  ultrastructure of, 12
  water content of, 54
Pseudohermaphrodites, nature of, 1157
Pseudo-parturition, nature of, 1141
Pseudo-pregnancy, nature of, 1113
Ptyalin, action of, 843
  speed of action of, 88
P.U., nature of, 1132
  placenta formation of, 1134
Puberty, female development at, 1115
  nature of, 1103
  spermatogenesis at, 1120
Pudic nerve, bladder emptying control by, 1024
Puerperium, events in, **1143**
  milk secretion in, **1146**
Pulmonary arterial pressure, magnitude of, **538**
Pulmonary artery, afferent impulses from, 593
  stretch receptors in, 767
Pulmonary circulation, 665–668
  exercise effect on, 673
Pulmonary veins, afferent impulses from, 593
Pulmonary ventilation, **716**
Pulmonary vessels, innervation of, 666
  pressures in, 665
Pulmonary pulse pressure, 666
Pulse, central and peripheral, **633**
  pressure, measurement of, **611**
  rate, average value of, 599
    exercise effect on, **675**
    reduced barometric pressure effect on, **776**
    respiration effect on, **600**
  reflected waves in, 634
  tracing, comparison of different arterial, **633**
  volume, measurement of, **636**
  wave, cause of, 629
    velocity of, 631
    velocity, measurement of, 635
Pulsus paradoxus, nature of, 567
Pupil, causes of constriction, 377
  cortical control of constriction of, 306
  diameter, depth of focus affected by, 393
  dilatation, causes of, 379
  innervation of, 289
  sympathetic control of, 227
Pupillary reflex, nature of, 289, **290**
Purine bases, excretion of, 899
Purines, composition of, 48
  metabolism of, 936
Purkinjé cells, cerebellum containing, **333, 334**
Purkinjé fibres, nature of, **553**
Purkinjé's figures, nature of, **412**
Purpura, causes of, 535
Pus cells, nature of, 524
Pus, formation of, 699
Putrescine, nature of, 97
Pyloric canal, movements of, 870
Pyloric sphincter, afferent nerve fibres from **353, 354**
  stomach emptying and, 872
Pyramidal cells, connections of, 295
Pyramidal tracts, composition of, 226
  effects of section of, 315
  origin and course of, 309–311
Pyrane, nature of, **34**
Pyranose, nature of, 33
Pyrexia, heat regulation in, 1052
Pyridine, blood platelet formation and, 527
Pyridoxol, transamination and, 924
Pyrimidine bases, structure of, **49**
Pyrimidine nucleus, nature of, **49**
Pyrimidines, metabolism of, 938
Pyrithiamine, nature of, 827
Pyruvic acid, alanine metabolism and, 927
  brain usage of, 746
  fat formation from carbohydrate and, 941
  fat metabolism and, 947
  heart usage of, 607, 746
  lactate formation from, 748
  oxidation of, **142, 143**
  structure of, **30**
Pyruvic carboxylase, vitamin B_1 action on, 826

Quantal summation, nature of, 130
Quinidine, heart affected by, 556
Quinine, skeletal muscle affected by, 138
QO_2, definition of, 745
  meaning of, 922

- **Rabbit**, ovulation in, 1106
- **Radial** pulse, jugular pulse and electrocardiogram with, **565**
  - record of, **634**
- **Radiation,** heat loss by, 1039
- **Radioactive** iodine, use of, 1073
- **Radioactive P,** blood volume measurement by use of, 495
  - metabolic studies with, 920
- **Radioactive** substance, circulation time measurement by, 623
- **Rami** communicantes, nature of, 343
- **Rationed** food, composition of, **837**
- **Reaction** of degeneration, nature of, 194
- **Reaction** time, measurement of, 318
- **Reactive** hyperæmia, nature of, 659
- **Rebound,** spinal reflexes showing, 260, **261**
- **Recessive** character, nature of, 1154
- **Reciprocal** innervation, eye muscles showing, 375
  - flexor reflex showing, **251,** 252
  - motor area stimulation and, 306–307
  - spinal reflexes and, 256–**257**
  - theory of mechanism of, 259–**260**
- **Recklinghausen** capsule, capillary pressure measurement by, 638
- **Recruitment,** extensor reflex showing, **261**
  - mechanism of, 262
- **Rectum,** defæcation and action of, 898
  - temperature in, **1035**
- **Red** cell count, nature of, 496
- **Red** nucleus, cerebellar connections with, 338
  - decerebrate rigidity and, 273
- **Reducing** enzymes, action of, 98
- **Reduction** division, nature of, **1097**
- **Reflex** action, characteristics of, 250
  - nature of, **212**
- **Reflex** arc, nature of, **213**
- **Reflex** contraction, frequency of nerve impulses in, 246
- **Reflex** figures, nature of, **232,** 233
- **Reflex** stance, factors concerned in, 277
- **Refraction,** factors concerned in, **385**
- **Refractory** period, cardiac muscle showing, 545
  - motoneurone showing, 241
  - nerve impulses and, 178
  - nerve showing, **179**
  - reflex response showing, **255**
  - skeletal muscle showing, 130
  - spinal flexor arc showing, 245–246
  - sympathetic ganglion showing, 238
- **Reinforcement,** conditioned reflexes showing, 326
  - spinal reflexes showing, 258
- **Reissner's** membrane, nature of, **454**
- **Remak's** ganglion, site of, 542
- **Renal** arterioles, urine flow regulation by tone of, 999
- **Renal** blood flow, glomerular filtration rate relation to, 999
  - measurement of, 1010
- **Renal** circulation, anatomy of, 991
- **Renal** function, history of, 980
- **Renal** glycosuria, nature of, 963
- **Renal** secretion, theories of, 994
- **Renal** tubules, function of, 1003
  - reabsorption from, 1004
  - secretion into, 1006
  - structure of, **991**
- **Renin,** action of, 1015
  - nature of, 661
  - site of production of, 991
- **Rennin,** action of, 862
- **Repetitive** discharge, motoneurones showing, 245
- **Reproduction,** history of, 1056
  - vitamin E effect on, 824
- **Reproductive** organs, hormonal influences on, 1103
- **Residual** air, measurement of, **717**
  - value of, **716**
- **Resonance** theory of hearing, 456
- **Respiration,** arterial pressure effect on, **766**
  - arterial pressure variation with, 667
  - artificial, 718
  - barometric pressure increase effect on, 778
  - bronchial movements in, 722
  - chemistry of, 722
  - diastolic filling of heart affected by, 567
  - glossopharyngeal nerve stimulation effect on, **858**
  - high altitude effect on, 775
  - history of, 490
  - lymph flow affected by, 691
  - mechanics of, 708
  - portal vein pressure affected by, 668
  - pulmonary pressures variation with, 666
  - pulse rate variation in, 600
  - reflex regulation of, 758
  - speech effect on, 471
  - of swallowing, nature of, 854, 859
  - vocal cords movement in, 468
- **Respiration** apparatus, use of, **785,** 791
- **Respiration** pump, types of, 718
- **Respiratory** centre, acetylcholine action on, 364
  - automaticity of, 753
  - carbon dioxide action on, 757
  - cardio-inhibitory centre affected by, **600**
  - location of, 749
- **Respiratory** exchange, exercise effect on, 804
  - measurement of, **785**
- **Respiratory** movements, chemical regulation of, 754
  - mechanics of, 708
  - methods of measurement of, **713**
  - regulation of, 749
- **Respiratory** quotient, calorific value of oxygen variation with, 792
  - definition of, 150
  - diabetes and, 967
  - exercise effect of, 153, 806, **943**
  - hibernation effect on, 944
  - measurement of, 788
  - nature of, 724
- **Restiform** body, afferent fibres in, 335
- **Rete** mucosum, structure of, **1027**
- "**Retention** of action current," nature of, 185–186
- **Reticular** cell, nature of, 525
- **Reticular** formation, motor cortex connections with, 311
- **Reticulocyte,** nature of, 519
- **Reticulo-endothelial** cells, red blood cell destruction by, 521
- **Reticulo-endothelial** system, function of, 525
  - jaundice and, 887
- **Reticulo-spinal** tract, function of, 227, **313**
- **Retina,** central connections of, 411
  - directional sensitivity of, **414**
  - electrical changes in, **406**
  - fatigue in, 429
  - lateral geniculate representation of, **288**
  - structure of, 402
- **Retinal** vessels, position of, **412**

Retinene, nature of, 406
  pigments formed from, 440
Retino-motor fibres, function of, 405
Retinoscopy, method of, 396
Retractor penis, action of, 1129
  innervation of, **162**
Retrograde degeneration, nature of, 224
Reverbatory circuits, cerebral cortex containing **298,** 299
Reverberation, speech affected by, 470
Reversion spectroscope, carboxyhæmoglobin determination by, 512
Rh factor, nature of, 503
Rheobase, nature of, 114
Rheotome, use of, **113**
Rhinomanometry, use of, 709
Rhodopsin. *See* Visual purple.
Riboflavin, actions of, 827
  hæmoglobin formation affected by, 521
  protein synthesis and, 923
Ribs, respiratory movements of, 711
Ricin, nature of, 701
Rickets, vitamin D relation to, 822
Righting reflexes, mechanism of, **280**
  nerve centres for, 282
Rigor mortis, factors affecting, 137
  lactic acid maximum in, 141
  muscle proteins in, 139
  nature of, 126
Ringer solutions, composition of, **62**
Ritter-Valli law, nature of, 181
"Robison ester," structure of, **34**
Rod, structure of, **404**
  theories of vision and, 437
  twilight vision and, 408
  vitamin A, effect on, 821
Room ventilation, standard conditions for, 779
Round ligaments, parturition and changes in, 1140
Rotation, reactions to, 283
Royal Air Force, respiratory tests for, 768
Rubber, gases soluble in, 81
Rubro-spinal tract, decerebrate rigidity and, 273
  function of, 227
Russell's tract, course of, **338**
Rutherford's theory of hearing, 459

Saccharin, circulation time measurement by, 623
Saccule, function of, 268, 271
Saliva, composition of, **843**
  functions of, 843
  secretion of, 844
Salivary "centre," location of, 845
Salivary duct, pressure in, 848
Salivary glands, changes during secretion in, 848
  conditioned stimulation of, 324
  innervation of, 846
  parasympathetic nerve supply to, 350
  structure of, **844**
  sympathetic nerve supply to, 347
Salivary secretion, electrical changes during, 851
  energy involved in, 852
  histological changes following, **850**
  thirst relation to, 901
  reflex arc of, **214**
Salmin, amino-acids in, 44
Salt, daily requirement of, **812**
  growth needs of, **789**
  intestinal absorption of, 906
  metabolic rate affected by lack of, 1064
  renal excretion of, 1010
  skin storage of, 1032
  sweat containing, 1029
Salt excretion, desoxycorticosterone effect on, 1062
Saponification, nature of, 27
Sarcolemma, nature of, 106
Sarcomere, nature of, **107**
Sarcoplasm, nature of, 106
Sarcoplasmatic membrane, nature of, 106
Saturation, difference threshold of, 424
Scala media, nature of, **454**
Scala tympani, nature of, **454**
Scala vestibuli, nature of, **454**
Schafer's method of artificial respiration, 719
Schiff-Sherrington phenomenon, nature of, 230
Sclero-proteins, nature of, 51
Scratch reflex, latent period of, 254
  nature of, **231**
  nociceptive stimulus inhibition of, **258**
Scopolamine, eye affected by, 391
Scurvy, capillary fragility in, 639
  vitamin C effect on, 828
Sebaceous glands, secretion of, 1028
  secretion, nature of, 27
Sebum, composition of, 1028
Secondary sexual characters, adrenal cortex effect on, 1063
"Second wind," nature of, 153, 769
Secretin, bile secretion affected by, 884
  nature of, 876
Segmenting movements, small intestine showing, **890**
Semen, composition of, 1123
Semi-circular canals, arrangement of, **268**
  function of, 271–272
Semi-lunar valves, heart response to damage to, 580
  structure of, 550
Seminal vesicle, androgens effect on, **1125**
  structure of, 1119
  sympathetic nerve stimulation effect on, 1129
Semi-permeable membrane, nature of, 57
Sensation, stimulus relationship to, 369
Sense organs, classification of, 368
Sense receptors, properties of, 219–**220**
Sensory nerve, vasodilator action of, **649**
Sensory paralysis, nature of, 316
Serum, bacteria affected by, 700
  calcium, protein affecting ionic state of, **814**
  colloid osmotic pressure of, 57
  composition of, 505
  definition of, 493
  phosphate, glucose absorption effect on, 907
  proteins of, 506
  protein, calcium ionisation relation to, **814**
    intestinal absorption of, 914
    osmotic pressure of, 71
  specific gravity of, 501
Serum albumin, amino-acids in, 44
  blood clotting affected by, 532
  iso-electric *p*H of, 74
  molecular weight of, 72
  properties of, 45, 506
Serum globulin, antibody formation and, 704
  blood clotting affected by, 530
  iso-electric *p*H of, 74

Serum globulin, molecular weight of, 45, 72
  properties of, 46, 507
Serine, structure of, 40
Sex, blood composition affected by, 500
  determination of, 1155
  hair growth affected by, 1032
  metabolic rate variation with, 832
  temperature difference in, 1035
  temperature regulation differences in, 1038
Sex chromosomes, nature of, **1156**
Sex hormones, specificity of, 1126
Sex-linked characters, nature of, 1156
Sham feeding, gastric secretion affected by, **860, 865**
Sham rage, hypothalamic stimulation causing, 358
  nature of, 233
Sheep, blood sugar value in, 954
Shingles, nature of, 235
Shivering, reflex nature of, 1054
  temperature regulation and, 1038
Shock, adrenal cortical hormones effect in, 1062
  events in, 663
  spinal transection in man causing, 234
  vascular changes in, 621
Shortening reaction, decerebrate preparation showing, 265
Short sight. *See* Myopia.
Simmond's disease, nature of, **1081**
Simultaneous contrast, nature of, 430
Singing, vocal cords action in, 468
Single-cone hypothesis of colour vision, 438
Sino-auricular node, structure of, 552
  vagus nerve stimulation effect on, 586
Sinus arrhythmia, nature of, 600
Sinus venosus, modification of, 552
Skatol, formation of, 930
Skeletal muscle, adrenaline action on, 1066
  aerobic metabolism of, 142
  all-or-none rule in, 127–**128**
  blood flow in resting, 629
  capillary area in, **742**
  capillary bed on exercise of, 675
  chemical composition of, 138
  chronaxie after nerve degeneration, 194
  chronaxie of, 114
  contraction, A.T.P. changes in, 148
    chemical changes in, 140
    chemical and heat changes in, 159
    mechanical efficiency of, 159
    mechanical energy of, 133–**134**
    pH changes in, 149
    physical changes associated with, 160
    physical changes in fibrils during, 107
    temperature effect on, 126
    theories of, **160**–161
  contractures in, 136–137
  creatine storage in, 933
  drugs acting on, 138
  electrical and mechanical responses of, 120, **131**
  enzyme systems in pyruvate formation in, **143**
  excitability of, 109
  exercise effect on blood flow in, 674
  fat content of, 943
  fatigue in, **135**–136
  glycogen storage in, 960
  heat production in, 155
  history of, 103
  initial length effect on contraction of, **124,** 125
Skeletal muscle, iodoacetate action on, 148
  isometric contraction of, **123–124**
  isotonic contraction of, **122**
  latent period of, **122–123**
  maximum work of, **134**–135
  myoglobin function in, 744
  neuro-muscular spindles in, **217–218**
  optimal rate of contraction of, 135
  oxygen usage of, 145
  potassium and sodium content in, 81–**82**
  propagation of contraction in, 127
  reaction of degeneration in, 194
  red, postural reflexes and, **279**
  refractory period of, 130
  relaxation, nervous connections effect on, 316
  starvation and loss in weight of, **797**
  tetanus of, 128–**129**
  tone, afferent impulses concerned in, 316
    cerebellum removal effect on, 339
    mechanism of, **131**
    sleep and, 329, 330
  thyroxine formation by, 1073
  twitch, recording of, **121–122**
  types of, 264
  vasodilatation, fainting due to, 656
  vasodilator nerves to, **647**
  viscosity, athletes and, 155
  voluntary contraction of, 132–**133**
Skin, absorption by, 1031
  blood flow to, 629
  blood supply at different temperatures to, 1044
  blood vessels, heat regulation effected by, 1043
    temperature sense relation to, 476
  clothing effect on heat loss by, 1046
  history of, 981
  reflex vaso-motor response to temperature change in, 1053
  sensations in, 474
    spinal cord tracts for, 224–**225**
  storage functions of, 1032
  structure of, **1027**
  temperature variation in, **1035**
    reflex responses to change in, 1053
  vascular reactions of, 662
  vessels, hæmorrhage effect on, 680
    vasodilator nerves to, 647
  water loss from, 811
Sleep, hypothalamic lesions causing, 358
  pupil size in, 378
  state of, 329
Sloth, muscle tonus in, 274
Small intestine, absorption from, 903
  digestion in, 875
  movements of, **890**
Smell, nervous pathways of, 290
  sense of, 485
Smell centre, location of, 305
Smooth muscle. *See* Plain muscle.
Sodium, cardiac muscle affected by, **581**
  daily requirement of, **812**
  movement to and from intestine of, 908
  red blood cell treatment of, 510
  skeletal muscle affected by, 138
Soft palate, deglutition and movement of, **855**
Sole, sweat secretion on, 1031
Solubility product, bone calcification and, 817
Somæsthetic area, functions of, 302
Sound, localisation of, 463
Sound pattern theory of hearing, 459

Specific dynamic action, cause of,803
  intimate nature of, 928
  protein showing, **916**
Specificity, conditioned reflexes showing, 325
Spectral colours, properties of, 372
Spectroscope test, colour blindness detection by, 435
Speech, acoustics of, 469
  neurology of, 470
Spermatid, nature of, 1098
Spermatocyte, nature of, 1098
Spermatogenesis, hormonal control of, 1120
  process of, 1098
Spermatozoa, fertilisation by, **1100**
  maturation of, 1099
  movement in female tract of, 1128
  numbers in semen of, 1123
  properties of, 1122
Spherical aberration, eye showing, 394
Sphincter of Oddi, function of, 885
  nature of, 882
Sphincter vesicæ, action of, 1017
Sphygmograph, use of, 632
Sphygmomanometer, use of, **612**
Spike-potential, nature of, **177**
Spinal animal, brain section producing, **273**
  properties of, 229
Spinal autonomic afferent fibres, path of, **352**
Spinal cord, acetylcholine synthesis in parts of, 364
  action potentials in, **246**, **247**
  afferent paths in, 224
  autonomic centres in, 355
  efferent paths in, 226
  micturition centres in, 1025
  nerve tracks in, 223, **225**
  section. *See* Spinal transection.
  structure of, 209, 222
  transmission in, 245
  trophic functions of, 234
Spinal cord canal, ciliary movement in, 169
Spinal cord lesions, micturition affected by, **1023**
Spinal flexor arc, refractory period of, 245–246
Spinal frog, properties of, 214
Spinal lemniscus, composition of, 284, **285**
Spinal nerve, composition of, **345**
Spinal reflex, algebraic summation of excitation and inhibition in, 259
  characteristics of, 250
  hypothalamic stimulation effect on, 357
  nature of, **230**
  vagus stimulation effect on, **359**
Spinal roots, action potentials in, 247, **248**
Spinal root section, effects of, 228
Spinal shock, nature of, 230
Spinal transection, blood pressure affected by, 642
  effects of, 227, 229
  effects in man of, 233
Spinal vasomotor centres, action of, 653
Spino-cerebellar tracts, composition of, **225**, 226
  destination of, 335–337
  functions of, 292
Spino-thalamic tract, composition of, 224
  pain relief by section of, 235
Spireme, nature of, 1096
Spirometer, lung volumes determined by, **717**
  respiratory exchange measurement by, **786**
Splanchnic nerve, composition of, 346, 646
Spleen, blood platelets affected by, 535
  circulation in, 681
Spleen, contractions of, **681**
  functions of, 682
  monocyte formation in, 525
  sympathetic stimulation effect on, 348
Splenectomy, red blood cell fragility affected by, 509
Spongin, nature of, 51
Spongioblasts, retina containing, **403**
Squid nerve, action potentials in, **176**
Stagnant anoxia, nature of, 771
Staircase effect, nature of, **545**
Standard deviation, definition of, 7
Stannius ligatures, nature of, 542
Stapedius, function of, 452
Starch, properties of, 37
  salivary digestion of, **843**
Starvation, bile secretion during, 885
  blood amino-acids in, **917**
  blood lipides affected by, 911
  fat usage in, 943
  glucose tolerance curve affected by, 972
  liver glycogen in, 957
  metabolic rate affected by food intake, 802
  metabolism in, 796
Static ataxy, nature of, 316
Static reflexes, types of, 276
Statistics, use of, 7
Stato-kinetic reflexes, types of, 282
"Steady state," nature of, 151
Stearic acid, nature of, 26
Stellate ganglion, composition of, 343
Stenopœic aperture, use of, 396
Stephan's law, nature of, 1039
Stepping reflex, negative successive induction in, 254
Stercobilin, nature of, 883
Stereognosis, cortical areas for, 303
Stereoscopic vision, factors in, **447**
Sterno-cutaneous muscle, strength-duration curves from, **115**
Sterols, metabolism of, 952
  structure of, 27–29
Stethograph, use of, **713**
Stilbœstrol, structure of, **1111**
Stomach, absorption from, 903
  alcohol action on, 839
  deglutition effect on, 858
  distension, gastric secretion from, 867
  emptying of, 871
  hæmopoietin formation in, 520
  hormone secretion by, 866
  movements of, 870
    vomiting effect on, 874
  nerve stimulation effect on movement of, **165**, 166
  parasympathetic stimulation effect on, 350
  pressure in, 874
  salivary digestion in, 844
  sympathetic stimulation effect on, 348
Strabismus, causes of, 375
  hypermetropia causing, 397
Stratum pigmenti, function of, 403
Strength-duration curve, *a* and *γ* components of, **115**
  nature of, **113**, 114
Stretch receptors, structure of, **217–218**
Stretch reflexes, characteristics of, 263–**264**
  muscle contractions in, **131**
String galvanometer, principle of, 118
Stroke volume, exercise effect on, **675**
Stromuhr, types of, **624**
Strophanthin, heart affected by, 579

Strychnine, motor area stimulation after, 307
  spinal cord affected by, 263
Sturin, amino-acids in, 44
Sub-arachnoid cisternæ, nature of, 693
Subcutaneous fat, composition of, 939
Subcutaneous tissue, gas pressures in, 743
Subliminal excitation, motoneurones showing, 243
Subliminal fringe, nature of, 253, **254**
Substrate, definition of, 89
Successive contrast, nature of, 430
Successive spinal induction, nature of, 254
Succinic acid, heart usage of, 747
  structure of, **31**
Succinic dehydrogenase, action of, **99**
Succus entericus, secretion of, 888
Suckling, uterine movements affected by, **1143,** 1148
Sucrose, properties of, 36
Sulphæmoglobin, formation of, 512
Sulphanilamide, *p*-aminobenzoic acid competition with, 820
Sulphate, urine containing, 984
Sulphonal, urine affected by, 987
Sulphur, amino-acids containing, 931
  enzymes containing, 90
  equilibrium, nature of, **785**
  metabolism of, 23, 931
Summation, central inhibitory state showing, 259
  conditioned reflexes showing, 325
  nerve and types of, 178
  reflex response showing, **253–254**
Summation of contractions, plain muscle showing, **164,** 166
  skeletal muscle showing, 128–**129**
Summation of stimuli, cardiac muscle showing, 545
  nature of, 115
  plain muscle showing, 166
  sympathetic ganglia showing, 239
Superior cervical ganglion ratio, of pre- and postganglionic fibres in, 349
  transmission in, 237
Superior colliculi, visual functions of, 289
Supplemental air, value of, **716**
Supporting reactions, nature of, 277
Suppressor bands, electrical activity of cortex affected by stimulation of, 321
Suppressor strips, location of, 308
  mode of action of, **314**
Supra-optic nucleus, function of, 358
Supra-opticohypophyseal tract, function of, **1089**
Surface tension, hydrotropy and, 909
  nature of, 66
Suspensoid, definition of, 70
Suspensory ligament, action in accommodation of, **390**
  action in eye of, 388
Swallowing, pulse rate affected by, 600
  respiration affected by, 767
Sweat, climate effect on output of, **1049**
  composition of, **1029**
Sweat glands, nerve supply to, 347
  structure of, 1029
Sweating, acclimatisation effect on, 1051
  factors affecting, **1030**
  skin temperature affected by, 1045
  temperature effect on, 1040
Sympathetic chain, composition of, 343
Sympathetic efferent paths, course of, 227
Sympathetic ganglia, acetylcholine formation in, **363**
  chemical transmission in, 361
  distribution of, 343
  transmission in, 237
Sympathetic nerve stimulation, adrenal medulla affected by, 1067
  gastric secretion with, 864
  heart affected by, **589**
  intestinal movements affected by, **893**
  lung volume affected by, **721**
  salivary glands affected by, 847, 851
Sympathetic nerves, distribution of, 346
  sweat glands supplied by, 1031
Sympathetic nervous system, afferent fibres of, **352**
  chemical transmission in, 360
  composition of, 343
  heart innervation by, **585**
  hypothalamic stimulation effect on, 357
  irradiation in, 359
  para-sympathetic comparison with, 351
Sympathetic plexuses, nature of, 346
Sympathetic vasoconstrictor nerves, tonic action of, 642
Sympathetic vasodilator nerves, 646
Sympathin, nature of, 167, 360
Synapse, conduction at, 236–**237**
  definition of, 195
  structure of, 205
Synapsis, nature of, 1099
Synaptic block, nature of, 207
Synaptic conduction, nature of, 236–**237**
Synaptic delay, nature of, 207
  sympathetic ganglion showing, 238
Synaptic knobs, nature of, 205, **206**
  neurone stimulation and, **242,** 243
Synaptic potential, nature of, 240
Synaptolemma, nature of, 205
Syngamy, nature of, 1099
Synovial fluid, composition of, 818
Systole, cardiac events in, 556
  pressure changes in heart in, **559**
Systolic murmur, cause of, 563
Systolic pressure, measurement of, **611**
  measurement in man of, **612**

Tachypnœa, nature of, 754
Tactile sensibility, test for, **477**
Talbot Plateau law, nature of, 422
Taste-buds, structure of, **484**
Taste centre, location of, 305
Taste, nervous pathways of, 290
  sense of, 483
Taurine, cysteine relation to, 931
  nature of, 883
Taurocholic acid, nature of, 883
Tear fluid, secretion of, **380,** 381
Tectorial membrane, function of, **454**
Tecto-spinal tract, function of, 227
Teeth, vitamin C effect on growth of, 829
  vitamin D effect on formation of, 823
Telephone theory of hearing, 459
Temperature, blood clotting affected by, 528, 534
  heart affected by changes in, 580
  nasal mucosa affected by changes in, **709**
  nerve conductivity affected by, 186
  nerve excitability affected by, 181
  oxygen dissociation curve affected by, 729
  plain muscle affected by, 169

Temperature, pulse rate affected by, 599
  red blood cell fragility affected by, 509
  renal secretion affected by, **997**
  sense of, 474
  skin temperature affected by, **1035**
  spermatozoa affected by, 1122
  sweating response to rise in, **1030**
Temperature regulation, nervous control of, 1053
Temperature sense, central nervous pathways for, **225**, 226, 284
  receptors for, 216, **217**
Temporal lobe, functions of, 300
Tendon receptors, function of, 218
Tendon reflexes, characteristics of, 264–265
Tendril fibres, origin of, 335
Tension receptors, nature of, 218
Tension-length relationship, resting and contracted muscle, 125–**126**
Tensor tympani, function of 451,
Test meal, nature of, 859
Testis, anterior pituitary effect on, 1125
  history of, 1056
  structure of, 1119
Testosterone, adrenal cortical hormones and action of, 1062
  structure of, 28, 1123
Tetanus, genesis of, 128–**129**
Tetanus toxin, species variation in susceptibility to, 701
Tetany, parathyroid glands and, 1076
Tetrahydro-β-naphthylamine, actions of, 1038
Thalamus, acetylcholine synthesis in, 364
  frontal lobe connections with, **301**
  functions of, 291
  hypothalamic connections with, 356
  motor cortex connections with, 311
Thalamus animal, brain section producing, **273**
Theobromine, structure of, 48
Thermal anhidrosis, nature of, 1052
Thermal interchange, temperature effect on, **1043**
Thermopile, use of, **156**
Thermostromuhr, principle of, 624
Thiamin, actions of, 826
Thiamine pyrophosphate, co-enzyme action of, **94**
Thiocyanate, formation and excretion of, 843
  thyroid gland affected by, 1076
Thiosulphate, excretion of, 1003
Thiouracil, thyroid gland affected by, 1075
Thiry-Vella fistula, nature of, 888
Thirst, cause of, 901
Thorax, respiratory movements of, 711
Threonine, structure of, 40
Threshold substance, nature of, 1004
Threshold values, method of, 367, 369
Thrombin, characters of, 529
Thrombocytes, nature of, 526
Thrombogen, nature of, 535
Thrombokinase, action of, 530
Thromboplastin, action of, 530
Thrombozym, nature of, 535
Thymine, structure of, **49**
Thymus gland, castration effect on, 1104
  function of, 1091
Thyro-arytenoid muscle, action of, 466
Thyroglobulin, nature of, 1072
Thyroid deficiency, effects of, 1069
Thyroid extract, diuretic action of, 1090
Thyroid feeding, effect of, **1071**
Thyroid gland, anterior pituitary interacti with, 1093
  blood flow to, 629
  carbohydrate metabolism affected by, 974
  functions of, 1069
  gonads interaction with, 1093
    relation to, 1104
  history of, 1055
  lactation affected by, 1145
  nerve supply to, 348
  temperature effect on, 1038
  thyrotropic hormone action on, 1085
Thyroid secretion, basal metabolic ra affected by, 795
Thyroidectomy, pituitary changes followin 1079
Thyrotropic hormone, action of, 1075, 1085
Thyroxine, actions of, 1073
  hair growth affected by, 1032
  milk secretion affected by, 1146
  structure of, 40, **1072**
  tyrosine relation to, 928
Tidal air, expired air composition affected 725
  value of, **716**
Tissue culture, method of, 12
Tissue fluid, factors affecting formation of, 6
Tissue regeneration, power of, 1153
Tissue respiration, factors concerned in, 74
  history of, 783
Tissue slice, measurement of metabolism 745
  metabolism of, **921**
Tissues, oxygen diffusion into, 741
T*m*, determination of, 1009
Tocopherol, structure of, **825**
Tongue, drugs action on, 484
  deglutition and action of, 854
Tortoise, heart anatomy in, 542
Touch, central nervous pathways for, 284
  cortical representation of, 303
  localisation of, 479
  reaction time of, 319
  receptors for, 216, **217**
    adaptation in, 220
  sense of, 476
  spinal cord tracts for 224–225
Touch corpuscles, location of, **1027**
Toxins, lymph absorption of, 692
  types of, 701
Toxophore group, nature of, 702
Trace conditioned reflex, nature of, 326
Trachea, air dust removal by, 779
Training, muscular performance affected 677
Transamination, nature of, 923
Transferring enzymes, action of, 101
  groups of, **96**
Transitional cells, characters of, 524
Transmethylation, nature of, 931
Transneuronal degeneration, nature of, 2
Transphosphatase, action of, 101
Traube-Hering waves, nature of, 652
Tricarboxylic acid cycle, fat oxidation an nature of, **143**
Tricuspid valve, structure of, 550
Trigeminal lemniscus, connections of, 28
Trioses, structure of, 30
Triple response, nature of, 662
Trochlear nerve, effect of injury to, 375
Trophic lesions, spinal cord section causin
Trophoblast, nature of, **1132**

Voluntary eye movements, control of, 376
Voluntary movement, cerebellar injury affecting, **341**
Voluntary muscular contraction, fatigue in, 136
  mechanism of, 132–**133**
Vomiting centre, action of, 875
Vomiting, intestinal secretion affected by, 888
  mechanism, of, 874
Vowels, frequency of components of, **469**
ulpian-Heidenhain-Sherrington phenomenon mechanism of, 136–137

Wallerian degeneration, nature of, 173
Warburg manometer, use of, **921**
War, diets affected by, 836
Wasserman reaction, nature of, 705
Water, absorption in intestine of, 905
  balance, pituitrin action on, 1088
  distribution in body of, 810
  diuresis, mechanism of, 1013
  large intestine absorption of, 895
  passage through membranes of, 79
  regulation, factors concerned in, 1090
  skin and lungs loss of, 1029
  skin storage of, 1032
‘ **Water-hammer** ” pulse, nature of, 629
Wave summation, nature of, 130
Weber's law, difference threshold of intensity in eye following, 423
  nature of, 371
  touch and, 478
Wedensky inhibition, nature of, 180, 196
ight-testing, Weber's law and, 370
eiss's laws, chronaxie and, **113**
ever and Bray effect, nature of, 455, 457
hey, nature of, 1149
hirling psychrometer, use of, 1047

White rami communicantes, composition of, 345
White light, composition of, 372
Widal test, nature of, 704
Wiggers manometer, use of, **557**, **558**

Xanthine oxidase, action of, **99**
Xanthine, structure of, **48**
X-chromosome, nature of, 1156
Xerophthalmia, vitamin A relation to, 821
X-ray, colon filling shown by, **897**
  deglutition movements followed by, 853
  fœtal circulation examined by, 1137
  intestinal movements shown by, **890**
  measurements on living eye by, **399**
  mutations induced by, 1159
  ovaries affected by, 1109
  renal circulation studied by, **993**
X-zone, nature of, 1059

Y-chromosome, nature of, 1156
Yeast, vitamin B_1 content of, 826
Yolk sac, blood cell formation in, **518**
Young's hypothesis of colour vision, 437

Z-disc, nature of, 107
Zein, amino-acids in, 44
  amino-acids lacking in, **808**
  properties of, 46
Zona pellucida, nature of, 1106
Zonula of Zinn. *See* Suspensory ligament.
Zwitterion, nature of, 76
Zymogen granules, nature of, 851
  pancreas containing, **878**

CENTRAL FOOD TECHNOLOGICAL RESEARCH INSTITUTE, * MYSORE. *
LIBRARY

PRINTED IN GREAT BRITAIN BY THE WHITEFRIARS PRESS LTD.
LONDON AND TONBRIDGE

14.5.97

C. F. T. R. I. LIBRARY, MYSORE-13

Acc. No. 368

Call No. L:3 N49

Please return this publication on or before the last DUE DATE stamped below to avoid incurring overdue charges.

To be issued from :

Due Date	*Return Date*	*Due Date*	*Return Date*
30.10.98	26.10.98		

. 368
3 . N49
NS.
of